

Proviral HIV-1 Hypermutation:
The Correlation of APOBEC3G/F and HIV-1 Vif in HIV-1 Disease Progression

by

Sujata Monika De

A thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
In partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Medical Microbiology and Infectious Diseases
University of Manitoba
Winnipeg, Manitoba

Copyright © 2011 by Sujata M. De

Abstract

APOBEC3 proteins, in particular APOBEC3G/F, are important innate host factors that contribute to protection from HIV-1 infection by inducing high levels of guanine to adenine nucleotide substitutions (termed hypermutation) during HIV-1 viral replication. These nucleotide substitutions occur at different rates and locations across the HIV-1 genome and are thought to be particularly more frequent in the pol region. The virus has evolved ways to counteract these host factors by inducing degradation of APOBEC3G/F proteins through protein interactions with HIV-1 Vif. The aim of this thesis is to characterize and investigate the role of APOBEC3G/F-mediated hypermutation in the HIV-1 genome.

We identified a subset of women from the Pumwani Commercial Sex Worker (CSW) cohort with significantly higher rates of hypermutated proviruses in pol. This degree of hypermutation correlated to less severe HIV disease progression as measured by CD4⁺ T cell count. This was in agreement with previous studies that evidence of APOBEC-mediated hypermutation correlate with reduced disease progression, confirming APOBEC3G/F proteins role in HIV-1 disease.

Furthermore, we investigated the *in vitro* and *ex vivo* interaction between HIV-1 Vif and APOBEC3G from subjects infected with hypermutated and non-hypermutated proviruses. *In vitro* studies indicated that HIV-1 Vif protein expression in subjects with hypermutated proviruses were quite divergent and levels of APOBEC3G also differed between subjects. *Ex vivo* studies in subjects with hypermutated proviruses indicated that endogenous APOBEC3G expression was greater than in subjects with hypermutated proviruses. Both studies suggest that host and viral factors such as APOBEC3G and HIV-

1 Vif are playing an influential role in HIV-1 pathogenesis. Further investigations into these interactions may lead to novel strategies into the development of therapeutic drugs for the fight against HIV-1.

Acknowledgements

I am grateful to many people who've supported me throughout my Master of Science degree at the University of Manitoba and please note the acknowledgements (in no particular order):

Dr. Frank Plummer, my supervisor, who has allowed me this incredible opportunity to work in his laboratory. It has been a great pleasure knowing that I will leave this lab more prepared for industry challenges, with a strong training background in HIV research.

Dr. Blake Ball, my co-supervisor, who has always been there, from day one – helping me plan and tackle all the difficulties of this project. You possess outstanding qualities as a person, those qualities then filter down to your students, your sense of understanding, the supportiveness, motivation and most notably your patience all make you a great (co-) supervisor. If the opportunity arose in the future it would be a pleasure to work with Blake again.

My committee members, Drs. Xiaojian Yao, Runtao He, and Sam Kung, which I feel there aren't enough words to significantly express my thanks for all their support, the encouragement, and scientific knowledge I have gained which has truly benefited me in this field, but beyond the scientific knowledge they have helped me better understand how to deal with troubleshooting not only in my work environment, but in my regular everyday life.

Dr. Ruey Chi-Su, my mentor, who has played a critical role in helping me analyze all the various aspects of my project that were troublesome. Ruey was always there to lend a helping hand and provide valuable advice over the last year of my project. Ruey is

one of the smartest scientists I know and her input and strategic advice went above and beyond what she needed to help me with, it was amazing to have the support she provided and thank you so much for being there in my time of need. Thank-you very much for your suggestions and also helping me revise this thesis on more than one occasion.

Dr. Adrienne Meyers, my mentor, my friend, someone that gave a large amount of help and advice on my project. Unfortunately there are too many things to individually list that Adrienne helped me with over the years, both as a supervisor, and friend. It is my greatest wish that you continue to embrace the work that you do, and continue conveying all the priceless knowledge you gave us to your students. Thank you!

Drs. Xiaojuan Mao and Ben Liang for their guidance into the world of sequencing and bioinformatics, this was a great piece to the puzzle for my continuation of learning. I appreciate all the time taken to sit with and pass along their understanding of this field. It's greatly attributed to my steep learning curve, and I truly feel that because of their great efforts I am a better and much more informed scientist.

Dr. Allison Land, what can I say, you introduced me to the world of APOBEC proteins. I want to express my great thanks and gratitude for everything you helped me with, thank you for always clarifying and answering my pesky questions through emails.

Leslie Slaney, Ian MacLean, Sue Ramdahin, Christine Mesa, Steve Wayne, Dr. Frauke Fehrmann, and Dr. Keith Fowke for their technical expertise and guidance; everyone from NML HIV group – a special thanks to Drs. Ma Luo, James Sainsbury and Adam Burgener, Thomas Bielawny, Rupert Capina, Jeff Tuff, and Paige Isabey for their assistance.

A special thanks to Angela Nelson, from the Department of Medical Microbiology

and Infectious Diseases, for all her help in administrative assistance.

Dr. Yao's laboratory (Dr. X. Yao, Zhujun, Yingfeng, Vincent, Kalleh, and Xiaoxia) for patiently helping me troubleshoot my experiments, borrowing reagents, and giving invaluable advice for my project. Thank-you all for your help!

I would like to thank everyone from the Plummer/Ball/Fowke HIV group, especially all the students (Winnie, Elnaz, Were, Aida, Nadine, Jill, Melissa, Derek, Meika, Caitlin, Jen, Catherine, Syeda, Lindsay, and Courtney) whom of all I have enjoyed my time here at this lab and each have helped me at one or more points through my degree.

Lastly, this work would not be possible without the participation of the women of the Pumwani CSW cohort and to all the staff members who dedicate their time and effort to pursue this valuable research cause.

Dedication

I dedicate this thesis to my parents who have always been patient, supportive, motivational, and inspirational to me through this incredibly long and painful journey. Both of you have sacrificed many things in your life for your children, and installed in us the value and importance of an education. I couldn't ask for any other parents in the entire world - you two are the most unselfish human beings who first and foremost put their children's needs first before their own. Mom, I would not be the woman I am today, without you. Thank-you both for telling me that anything is possible if you try and I love you both dearly.

To Jeff, I am very grateful to have such a wonderful, kind-hearted man in my life and words cannot even begin to describe how inspiring you are to me. You have been such a caring, loving, and supportive boyfriend (even from a distance) from day one and thank-you for always being patient through my good and bad days. You are my future, my heart, my soul, and my best friend. I love you.

As months and years have passed, I have found many difficulties with research and I have only learned by the mistakes that I've done, that my appreciation for virology has increased, particularly the field of HIV research. I have gained valuable knowledge for what research truly is, and the amount of dedication it takes to conduct research.

Table of Contents

Abstract	II
Acknowledgements	IV
Dedication	VII
List of copyrighted material for which permission was obtained	XVI
1.0 Introduction	1
1.1 Global impact of HIV-1/AIDS epidemic	1
1.1.1 Progression to AIDS	1
1.1.2 The urgency for a vaccine	3
1.2 Routes of HIV-1 transmission	4
1.2.1 HIV-1 pathogenesis	4
1.2.2 Clinical stages of HIV-1 infection	5
1.2.3 Treatment	7
1.3 HIV virus	9
1.3.1 HIV classification	9
1.3.2 HIV-1 structure and genome	10
1.3.3 Organization of the untranslated signaling regions of the HIV-1 RNA genome	11
1.3.4 HIV-1 Gag	16
1.3.5 HIV-1 Pol	16
1.3.6 HIV-1 Env	17
1.3.7 HIV-1 accessory/regulatory proteins	17
1.3.8 HIV-1 Vif	18
1.3.9 HIV-1 entry and reverse transcription	20
1.3.10 HIV-1 genome replication	21
1.3.11 Viral diversity in HIV-1	24
1.4 HIV-1 host restriction factors	24
1.4.1 AID / APOBEC family	25
1.4.2 APOBEC3G/F in HIV-1 infection	27
1.4.3 Mechanism of HIV-1 Vif inhibition of APOBEC3G/F	29
1.4.4 APOBEC3G/F-mediated hypermutation of the HIV-1 genome	31
1.5 The significance of CD4 ⁺ T cell and viral load counts in HIV-1 disease progression	32

1.6	APOBEC and disease progression	35
2.0	Rationale, hypothesis and objectives.....	36
3.0	Materials and methods	38
3.1	General laboratory materials	38
3.1.1.	Laboratory prepared solutions	38
3.1.2	Reagents.....	41
3.1.3.	Commercial kits.....	43
3.1.4	Specialized equipment	45
3.1.5	Software programs.....	45
3.2	Study population	46
3.2.1	Pumwani CSW cohort	46
3.2.2	Subject selection for APOBEC3G degradation.....	46
3.2.3	Subject selection for sequencing: examination of HIV-1 proviral hypermutation at the <i>pol</i> gene.....	47
3.3	Protocols.....	47
3.3.1.	Extraction of total genomic DNA.....	47
3.3.2	Amplification of HIV-1 provirus.....	47
3.3.2.1	Full length <i>pol</i> amplification.....	47
3.3.2.2	Partial <i>vpu/env</i> region amplification.....	51
3.3.3	Visualizing PCR products by gel electrophoresis	51
3.3.4	PCR product purification.....	53
3.3.5	DNA sequencing.....	53
3.3.5.1	Full length <i>pol</i> sequencing	53
3.3.5.2	Partial HIV-1 proviral genome sequencing	54
3.3.6	HIV-1 sequences and phylogenetic analysis	54
3.3.7	Detection of hypermutation at HIV-1 proviral genome	56
3.3.8	Statistics used for data analysis	56
3.3.9	Cell culture.....	56
3.3.10	Plasmids.....	56
3.3.11	Plasmid DNA amplification, purification and quantification.....	59
3.3.12	Restriction enzyme digests of vectors	61
3.3.13	Transfections of 293 T cells	61
3.3.14	Preparation of cell lysates.....	62

3.3.15	Quantification of protein concentration in cell lysates	63
3.3.16	Western blots	64
4.0	Results	65
4.1	HIV-1 proviral <i>pol</i> gene sequence reflects APOBEC-mediated hypermutation.	65
4.1.1	Defining hypermutated sequences and HIV-1 subtype assignment	65
4.1.2	Sequencing data reveal APOBEC3G/F-mediated hypermutation at HIV-1 proviral <i>pol</i> gene	71
4.1.3	APOBEC3G/F-mediated hypermutation at HIV-1 <i>vpu/env</i> region is	80
	independent of the hypermutation at the HIV-1 <i>pol</i> region and <i>vice versa</i>	80
4.2	The presence of HIV-1 hypermutation at proviral <i>pol</i> gene correlates with higher CD4 ⁺ T cell counts in HIV-1 ⁺ Kenyan CSW	85
4.2.1	Hypermutation of proviral <i>pol</i> sequences introduces premature stop codons.	85
4.2.2	Proviral <i>pol</i> sequences correlate with CD4 ⁺ T cell counts	88
4.3	HIV-1 ⁺ women with hypermutation at the proviral HIV-1 genome (<i>env</i> , <i>pol</i> , and <i>vpu</i>) have increased APOBEC3G expression.....	92
4.3.1	Rationale	92
4.3.2	<i>Ex vivo</i> expression of HIV-1 Vif and endogenous APOBEC3G proteins in PBMCs of HIV-1 ⁺ women with APOBEC-mediated <i>vpu/env/pol</i> hypermutation.....	94
4.4	<i>In vitro</i> expression of HIV-1 Vif in subjects with <i>vpu/env</i> hypermutation was reduced compared to non-hypermutated subjects.....	99
4.4.1	Rationale	99
4.4.2	Verification of plasmid construct by restriction enzyme digestion	100
4.4.3	Co-expression of wt APOBEC3G with HIV-1 Vif, isolated from subjects with HIV-1 <i>vpu/env</i> proviral hypermutation-mediated by APOBEC: <i>vif</i> isolated from subjects with or without hypermutation had similar effects on wt APOBEC3G expression	104
5.0	Discussion	108
5.1	Sequencing data suggest that hypermutation at HIV-1 proviral <i>pol</i> region is mediated by APOBEC3G/F: the significance.....	109
5.2	The presence of hypermutation at HIV-1 proviral <i>pol</i> region correlates with higher CD4 ⁺ T cell counts: the significance in disease progression	112
5.3	The presence of hypermutation at HIV-1 proviral <i>vpu/env</i> region is independent of that at the <i>pol</i> region: possible mechanisms and significance	114
5.4	Increased host APOBEC3G expression in <i>ex vivo</i> PBMCs of HIV ⁺ women: Vif-mediated regulation.....	116
5.5	The relationship between the level of HIV-1 Vif expression and the level of host	

APOBEC3G: Vif-mediated regulation	121
6.0 Conclusions and future directions	126
7.0 References	128
8.0 Abbreviations used.....	143

List of tables

Table 1: Pol primers used for nested-PCR, their gene locations, melting and annealing temperatures	49
Table 2: Thermocycler conditions for nested-PCR amplification	50
Table 3: Vpu/env primers used for nested-PCR, their gene locations, melting and annealing temperatures	52
Table 4: Sequencing primers for pol (full-length and partial) and their gene locations, melting and annealing temperatures	55
Table 5: HIV-1 subtype classifications from twenty-eight full length pol sequences generated in this thesis	69
Table 6: Sixteen samples in the pol region (nucleotide positions 4150-4919) demonstrated significant hypermutated HIV-1 provirus	72
Table 7: Composition of one hundred thirty-four samples infected with HIV-1 subtype A1 with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)	75
Table 8: Composition of twelve samples infected with HIV-1 subtype C with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)	76
Table 9: Composition of forty samples infected with HIV-1 subtype D with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)	77
Table 10: Composition of three samples infected with HIV-1 subtype G with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide	

positions 4150-4919)	78
Table 11: Composition of eight samples infected with HIV-1 recombinants with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)	79
Table 12: List of subjects where two regions of the HIV-1 genome was examined for APOBEC3G/F-mediated hypermutation	81
Table 13: Comparison of hypermutated subjects at two different regions of the proviral HIV-1 genome	82
Table 14: Three subjects that showed HIV-1 proviral APOBEC3G/F-mediated hypermutation at both vpu/env and at pol regions	84
Table 15: Number of premature stop codons present throughout the <i>pol</i> gene from sixteen samples with hypermutated pol proviruses	86
Table 16: PBMC list of subjects	93
Table 17: The effect of endogenous Vif protein expression of subjects on degradation of endogenous APOBEC3G protein	98
Table 18: Patient list and proviral HIV-1 hypermutation characteristics at vpu/env and vif	102
Table 19: The effect of Vif protein expression of subjects on degradation of wt APOBEC3G protein.....	107

List of figures

Figure 1: A global view of HIV infection in 2009	2
Figure 2: Process of HIV-1 disease progression in the absence of antiretroviral therapy over several weeks to years.....	6
Figure 3: Schematic of HIV-1 structure.....	12
Figure 4: Genomic organization of HIV-1 and the processes of transcription and translation of HIV-1 genes.....	13
Figure 5: Important sequences required for HIV-1 replication	15
Figure 6: Highlights of HIV-1 reverse transcription in the process of strand synthesis....	22
Figure 7: Molecular mechanisms of APOBEC proteins and Vif can influence the retroviral lifecycle	28
Figure 8: The protein interactions between Vif and APOBEC3G (A3G) domains.....	30
Figure 9: Two distinctive twin gradients are observed during editing of the HIV-1 DNA genome by activity of APOBEC3 proteins	33
Figure 10: Schematic of SVCMV-IN expression vector encoding several markers	57
Figure 11: Schematic of pCMV-HA mammalian expression vectors encoding several markers.....	58
Figure 12: Plasmid DNA yields from four different protocols.....	60
Figure 13: Relative proportions of HIV-1 subtypes and recombinants for 241 HIV-1 proviral pol sequences. (A) The percentage of the 241 pol sequences organized into a pie chart with the appropriate subtd recombinants are in pol sequence pool.	67
Figure 14: Neighbor-joining tree of 28 full-length pol sequences generated in this thesis	

along with all publically available pol HIV-1 reference sequences.....	68
Figure 15: Graphical representation of mutations detected by using the tool hypermut 2.0	70
Figure 16: Histogram depicting dinucleotide context of G to A hypermutation from proviruses in sixteen subjects.....	74
Figure 17: Nucleotide and protein coding of sixteen samples with significantly hypermutated proviruses at HIV-1 pol.....	87
Figure 18: Association of higher CD4 ⁺ count with hypermutated proviral DNA sequences in pol. CD4 ⁺ counts are compared between subjects with hypermutated proviral DNA sequences and subjects without hypermutated proviral DNA sequences	89
Figure 19: Correlation of adenine proportion in the pol region of proviral HIV-1 with CD4 ⁺ T cell counts from 235 subjects.....	90
Figure 20: APOBEC3G and Vif expression in PBMC cell lysates	95
Figure 21: Comparison of <i>vif</i> sequences with subjects that have hypermutated and non- hypermutated HIV-1 proviral <i>vpu/env</i> sequences	101
Figure 22: 1% agarose gel of <i>pst</i> I and <i>xba</i> I digest of expression vectors containing the <i>vif</i> and APOBEC3G insert.....	103
Figure 23: Expression of wt Vif, subject Vifs, and APOBEC3G proteins	105

List of copyrighted material for which permission was obtained

Figure 1: A global view of HIV infection in 2009 (UNAIDS) (page 2)
((192) Permission received on March 28, 2011.)

Figure 2: Hallmarks of HIV infection in the absence of antiretroviral therapy over several weeks to years (page 6)
((129) Permission received on January 13, 2011.)

Figure 6: Highlights of HIV-1 reverse transcription in the process of strand synthesis (page 23)
((57) Permission obtained January 17, 2010.)

Figure 7: Molecular mechanisms of APOBEC proteins and Vif can influence the retroviral lifecycle (page 29)
((78) Permission obtained January 9, 2011.)

Figure 8: The protein interactions between Vif and APOBEC3G (A3G) domains (page 31)
((17) Permission obtained February 4, 2011.)

Figure 9: Two distinctive twin gradients are observed during editing of the HIV-1 DNA genome by activity of APOBEC3 proteins (page 34)
((189) Permission obtained February 2, 2011.)

Figure 21: Comparison of *vif* sequences from subjects with hypermutated and non-hypermutated proviral *vpu/env* sequences (page 102)
((117) Permission obtained February 2, 2011.)

Table 18: Patient list and proviral HIV-1 hypermutation characteristics at *vpu/env* and *vif* (page 103)
((117) Permission obtained February 2, 2011.)

1.0 Introduction

1.1 Global impact of HIV-1/AIDS epidemic

Significant progress has been made over the last several years in the treatment and understanding of the infection and pathogenesis of human immunodeficiency virus (HIV) type 1 (HIV-1), the causative agent of acquired immunodeficiency syndrome (AIDS). The virus currently affects the lives of an estimated 33.4 million individuals worldwide, as stated by the 2010 Joint United Nations Programmes on HIV/AIDS (UNAIDS) (192). Sixty million people have been infected since the beginning of the epidemic (108), and approximately 25-28 million deaths are due to HIV related-causes (183, 192). The region of the world most affected by this epidemic has been sub-Saharan Africa; in 2009 this is where 71% of all new HIV infections were found (Figure 1) (3, 30, 108, 153, 192).

1.1.1 Progression to AIDS

During acute HIV-1 infection, a significant drop in cluster of differentiation 4 (CD4)⁺ T cell counts is prominent, followed by an incomplete rebound in CD4⁺ T cell numbers leading to the latency period, where homeostasis is reached. Gradually, CD4⁺ T cell numbers decline over several years until CD4⁺ T cell depletion reaches a critical threshold of two hundred cells per millilitre (mL); this is considered the onset of AIDS (63, 144, 168, 177). The host immune response is so severely compromised that a large variety of opportunistic infections may arise. HIV-1 associated opportunistic infections such as pneumonia, caused by the fungus *Pneumocystis jirovecii*, Kaposi's sarcoma, and cytomegalovirus encephalitis are hallmarks of progression to AIDS (21, 175, 198).

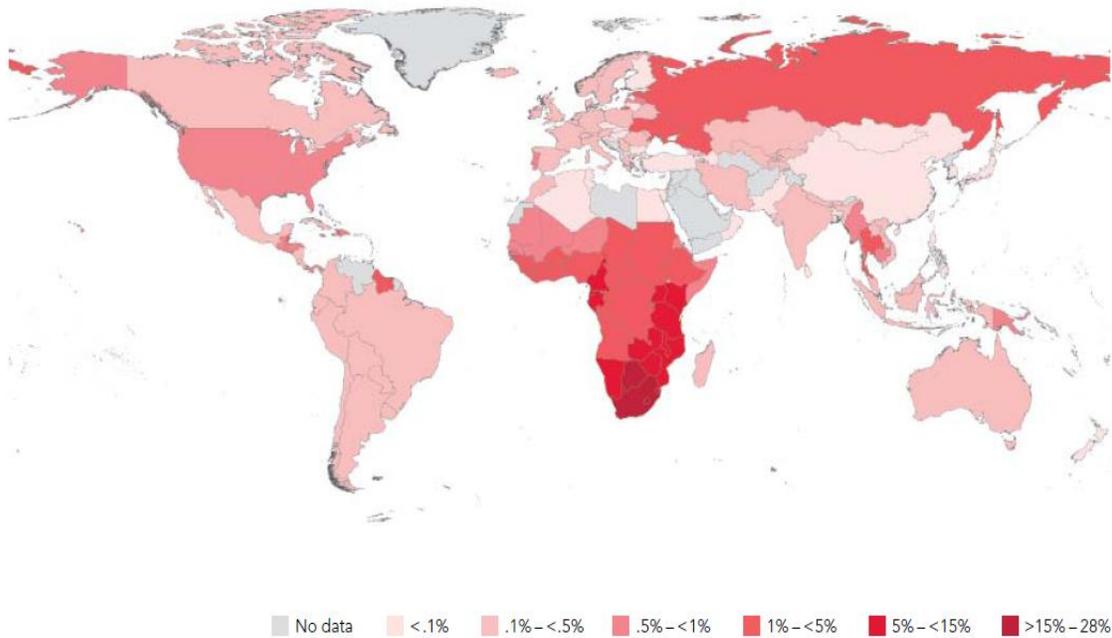


Figure 1: A global view of HIV infection in 2009 (192). The geographic areas with the highest infection rates of HIV are shown in darker shades of red; the geographic areas with the lowest infection rates of HIV are represented by lighter shades of red. The most affected region with the highest HIV infection rates is sub-Saharan Africa. Figure reproduced with permission from the UNAIDS report on the global AIDS epidemic 2010. Permission obtained March 28, 2011.

1.1.2 The urgency for a vaccine

The two main approaches utilized thus far towards development of a HIV vaccine include the development of antibody responses targeting HIV-1, viral envelope proteins and inducing the host CD4⁺ and CD8⁺ T cells to elicit strong immune responses (132). Unfortunately, no effective or approved preventative vaccine yet exists (197). There have been several unsuccessful clinical trials for vaccines such as the STEP trial by Merck, (15, 125, 132, 193), but none have provided protection until recently (the Thai RV144 Phase III trial showed a moderate 30% reduction in infections) (193).

Targeting HIV-1 remains problematic as there are several biological obstacles in developing a protective vaccine. These include: viral escape from host immune responses, the persistent infection of a cell due to HIV-1 incorporation into the host cell's genome, the lack of suitable animal models, the infection by HIV-1 of the very immune cells and exploitation of the immune activation pathways required for successful vaccination for its own replication, and most importantly the profound genetic variability of HIV-1 (66, 183). With up to 30% amino acid sequence differences between HIV-1 subtypes, an effective vaccine must be broadly active to prevent emergence of resistant viruses (66, 183).

The creation of a protective HIV vaccine is difficult as there are several variables that must be considered. However, alternative approaches to preventative strategies and therapeutics in controlling HIV infection are a necessary and complimentary approach of HIV.

1.2 Routes of HIV-1 transmission

Despite the initial misconception that HIV-1 transmission occurs mainly through homosexual contact, it is now well established that HIV-1 infection is primarily spread heterosexually through mucosal cell surfaces in the urogenital tract (76, 101, 138). Mother-to-child HIV-1 transmission can also occur in utero, during labour and delivery or through breastfeeding, which is the number one cause of HIV-1 acquisition by children (14, 38, 99). Sharing needles among intravenous drug users is another main route of HIV-1 transmission, as those individuals comprise an estimated 3 million people living with HIV-1 (192). The sexual transmission efficiency of HIV-1 is significantly influenced by bacterial, viral and sexually transmitted infections that are present at the time of the exposure (28, 73, 74, 139, 141); other transmission factors also include the concentration of the virus (viral load) and also the viral subtype one is exposed to (29, 30). It is important to note that the concentration of the virus in bodily fluids (blood, saliva, semen, vaginal fluid) and at what clinical stage of HIV-1 infection an infected individual is at can affect HIV-1 transmissibility (152).

1.2.1 HIV-1 pathogenesis

The mechanisms underlying HIV-1 pathogenesis are complex as they involve a combination of host and viral factors (71). Age, genetics, co-infection with another virus or microbe, and viral virulence will influence the pathogenesis of HIV-1. These factors determine the rate and severity of disease progression in an infected individual (13, 31, 36, 136). HIV-1 targets and generally replicates in a number of cell types such as

monocytes, macrophages, dendritic cells, and the most important reservoir, CD4⁺ T cells (169, 184). All these cell types express a common receptor, CD4, which make them susceptible to HIV-1 infection. The infection and eventual destruction of these cell types leads to CD4⁺ T cell depletion; the cause of HIV pathogenesis by weakening the host's immune system and essentially leaving the host unable to fight off infections caused by other pathogens.

1.2.2 Clinical stages of HIV-1 infection

The clinical stages of HIV-1 infection in individuals not on anti-HIV therapy can be split into three distinct phases: the acute phase, the clinically latent or asymptomatic phase, and eventual progression to AIDS (Figure 2) (129, 147, 158). The acute phase follows initial infection with HIV-1 and is characterized by a rapid depletion of CD4⁺ T cells, large amounts of virus production and induction of anti-HIV CD8⁺ and antibody responses (71, 129). It is thought that the cell-mediated immune response effectively reduces and controls viral load initiating the clinical latency phase (81, 134, 190). No symptoms are apparent in this clinically-latent phase of viral suppression which is characterized by a relatively low viral load and gradual slow depletion of CD4⁺ T cell count can last for a period of months to years (3, 129, 166). When CD4⁺ T cell counts drop below a critical point of 200 CD4⁺ T cells/mL of blood as opposed to the normal count of about 1000 CD4⁺ T cells/mL of blood, clinical AIDS is diagnosed (131, 166). Individuals with AIDS are unable to mount an effective immune response and are vulnerable to opportunistic infections and malignancies. Along with the decrease in CD4⁺ T cells in the blood and susceptibility to a variety of fatal opportunistic infections, AIDS

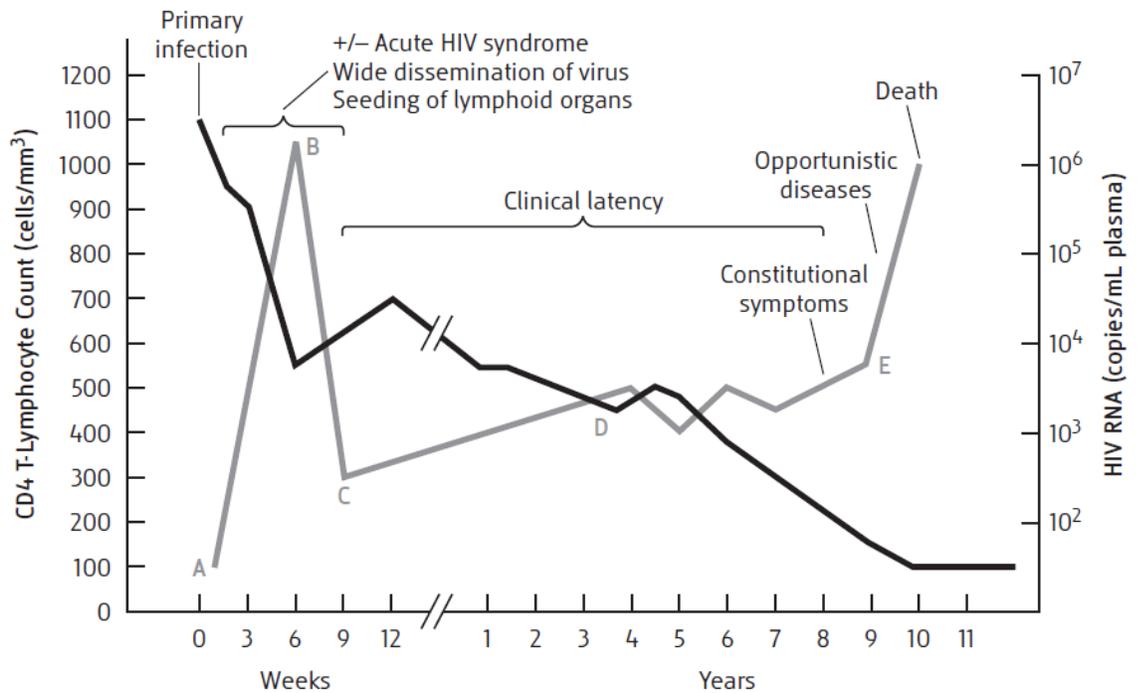


Figure 2: Process of HIV-1 disease progression in the absence of antiretroviral therapy over several weeks to years (129). The course of HIV-1 infection occurs in three clinical phases (acute, clinical (asymptomatic), and AIDS phase). The primary infection of HIV-1 viral load is lowest in the first few weeks of infection (point A) and viral load rapidly increases to its highest at 6 weeks (point B). A decrease in viral load occurs and is maintained over several years (point C, D). Viral load gradually increases after several years when constitutional symptoms and opportunistic diseases arise (point E). CD4 T-lymphocyte counts are high before primary infection of HIV-1 and rapidly decrease in the acute phase, rebound partly when the viral load is controlled. Essentially over several weeks to years CD4 T-lymphocytes decrease slowly in the asymptomatic phase where it leads to levels below 200 cells/mm³ (considered to be the onset of AIDS) as viral loads progressively increase. Figure reproduced with permission from Sara Teitelman from the Elizabeth Glaser Pediatric AIDS Foundation on January 13, 2011.

is characterized by other factors such as antigen presentation dysregulation, decrease in HIV-1 specific CD8⁺ CTL responses, decrease in HIV-1 specific antibodies concurrent with generalized hypergammaglobulinemia, and sharp increase in viral load (4, 43, 111)

1.2.3 Treatment

HIV drug therapies have extended the quality of life and life expectancy of those infected with HIV. This is due to the evolution of HIV treatment and therapy options available today (27, 116, 138). Currently, there are twenty-five approved antiretroviral drugs on the market for the treatment of HIV/AIDS patients (41, 42). The use of highly active antiretroviral therapy (HAART) in developed countries has had an enormous impact on AIDS patients and it is the method of choice to treat HIV-1 infected individuals. Presently in developing countries, HAART treatment is uneconomical to large-scale implementation because of the costs and the difficulties associated with administering treatment (108). However, external agencies such as the United States President's Emergency Plan for AIDS Relief (PEPFAR) and UNAIDS have helped millions of individuals on HAART worldwide (150, 192). There are other treatment options available, such as gene therapies, that are being explored for the treatment of HIV-1 infected individuals, however their use is prohibited due to their toxicity (163, 182).

There are four major groups of antiretroviral drugs that target and/or halt different parts of the HIV-1 replication cycle: non-nucleoside reverse transcriptase inhibitors (NNRTIs), nucleoside analogue reverse transcriptase inhibitors (NRTIs), protease inhibitors (PIs), and fusion inhibitors (41, 116). NRTIs and NNRTIs, work by inhibiting

the action of reverse transcriptase (RT) and therefore prevent the formation of proviral deoxyribonucleic acid (DNA) (37, 40, 41). The first NRTI to be licensed in the United States for the treatment of AIDS was Zidovudine (AZT) (37, 42). NNRTIs, unlike NRTIs, directly inhibit RT by binding in a reversible and non-competitive manner to the enzyme RT (37, 40, 100). Of the twenty-five drugs available, there are currently ten drugs classified under PIs which work by inhibiting the action of protease (PR), an enzyme utilized by HIV-1 to replicate (41). Enfuvirtide, a fusion inhibitor, is the only available drug in this class that interacts with HIV-1 in a specific manner by coiling with the heptad region (HR) region of the viral glycoprotein, gp41 (41).

HAART involves a combination (two or more) of NRTIs as well as PIs that act synergistically to decrease the viral load which determines the effectiveness of the treatment (42). Depending on the severity of infection, there are three types of combination regimes an infected individual can be treated with: PI with 2 NRTIs, NNRTI with 2 NRTIs, and 3 NRTIs (210). HAART combination regimes reduce viral load and usually lead to a gradual increase of CD4⁺ T cell counts in patients and in few cases CD4⁺ T cells counts are restored back to normal levels (210). HAART has increased the quality of life by minimizing the occurrence of opportunistic infections and it has dramatically affected the short-term mortality of HIV-1 infected individuals by reducing the levels of circulating virus (138, 144, 183).

However, the development of new anti-HIV treatments is necessary as many individuals experience side effects to current treatments and increasing rates of viral resistance. Furthermore, anti-HIV drugs are expensive to produce and deliver. Therefore it's important to develop new and more cost-efficient treatment options. Further

understanding of natural factors capable inhibiting HIV replication would aid in these efforts.

1.3 HIV virus

HIV is an enveloped virus encapsulating two copies of single-stranded (ss) positive ribonucleic acid (RNA). It is estimated to be one hundred nanometer (nm) in diameter, originates from the genus *Lentivirus*, and belongs to the family *Retroviridae*. Other members of this family include: simian immunodeficiency virus (SIV), equine infectious anemia virus, feline leukemia virus, and human T-lymphotropic virus (20, 58). The *Retroviridae* family is unique in that no other family of viruses are able to reverse-transcribe its genome from RNA to a DNA intermediate (using a reverse transcriptase enzyme) for replication; whereas other RNA viruses (double-stranded (ds), ss positive sense, ss-negative sense) transcribe their genome via a RNA intermediate (using a RNA-dependent RNA polymerase) in the cytoplasm of the infected cell.

1.3.1 HIV classification

There are two types of HIV, HIV-1 and HIV-2 (30, 86), which have arisen from zoonotic transmission from two different ancestral strains of SIV (65, 171). The predominant virus circulating worldwide is HIV-1 while HIV-2, which appears to be less transmissible and less pathogenic, is concentrated primarily in West Africa. Based on phylogenetic analysis, different HIV-1 isolates can be further divided into four groups: Major (M), Outlier (O), Non-M and Non-O (N), and P group (148, 154). On a global scale, M group is responsible for the majority of HIV-1 infections while the O- and N-

group are less frequently observed (183). Isolates belonging to the M group are further subdivided into several subtypes based on further phylogenetic analysis of their sequences: A, B, C, D, F, G, H, J, K, and circulating recombinant forms (CRFs) (30, 126, 183). Subtypes and CRFs are found based on geographical locations where subtype C is the most highly distributed worldwide, while subtype B is the most prevalent in the Western world (7, 44, 89). The overall genomic organization of HIV-1 subtypes is similar, but sequence diversity can range between 5-35% in different genes between subtypes (89, 103). The emergence of HIV-1 subtypes and their biological relevance is of considerable interest as many groups around the world are conducting studies on subtype diversification to determine if a relationship to disease progression exists (7, 10, 49, 103, 104, 109). However, correlation between specific subtypes and HIV-1 disease progression is controversial as some groups have not found any associations (62, 86).

1.3.2 HIV-1 structure and genome

The HIV-1 viral genome is a dimer of linear, positive-sense, ss RNA. Each RNA molecule is 9.2 kb in length and encodes nine genes which include three structural proteins: group specific antigen (*gag*), polymerase (*pol*), and envelope (*env*) as well as non-structural proteins (52, 174, 202). The non-structural proteins can be further subdivided into regulatory proteins: trans-activator of transcription (*tat*), differential regulator of expression (*rev*) and accessory proteins: viral infectivity factor (*vif*), negative regulator factor (*nef*), virus protein U (*vpu*), and viral protein R (*vpr*) (45, 167, 202).

There are long-terminal repeats (LTR) common to all retroviruses located at each end of the RNA molecule. The HIV-1 genome is flanked by a LTR at the 5' and 3'

location, essentially lengthening the genome in size to 9.7 kb. The 5' and 3' are essential for the virus to successfully integrate into the host genome (135, 213). The expression of integrated HIV-1 DNA is strongly regulated by sequences in the transcriptional control region of the viral LTR (containing the promoter of HIV-1, the transcriptional start site, and enhancer regions) and this is recognized by the host cell transcriptional machinery (163).

The RNA molecules are surrounded by capsid proteins: p17 (matrix), p24 (capsid), p7 (nucleocapsid), and p6 proteins (48, 56, 59). These proteins are encoded by the 55 kilodalton (kDa) *gag* precursor, which will be described in a later section. Integrase (IN), PR, and RT are tightly bound to the RNA molecules within the capsid. The structure of HIV-1 and the viral genome (along with their transcription and translational processes of genes) are represented in Figure 3 and Figure 4, respectively.

1.3.3 Organization of the untranslated signaling regions of the HIV-1 RNA genome

The viral genome is transcribed by normal host transcriptional machinery and it has many features of normal messenger RNA (mRNA). The RNA is capped at the 5' end, using a common methylguanosine cap and it has a string of two hundred nucleotide long polyadenylation (Poly (A)) sequences located at the 3' end (163). The implication of the HIV-1 RNA genome being very similar to host mRNA means the viral genome easily transcribes into the host genomic material.

There are several other key sequences present in the LTR of the HIV-1 genome such as the short sequence which is required for viral replication and integration. This short sequence consists of a repeated (R) region that is located at the 5' end of the RNA

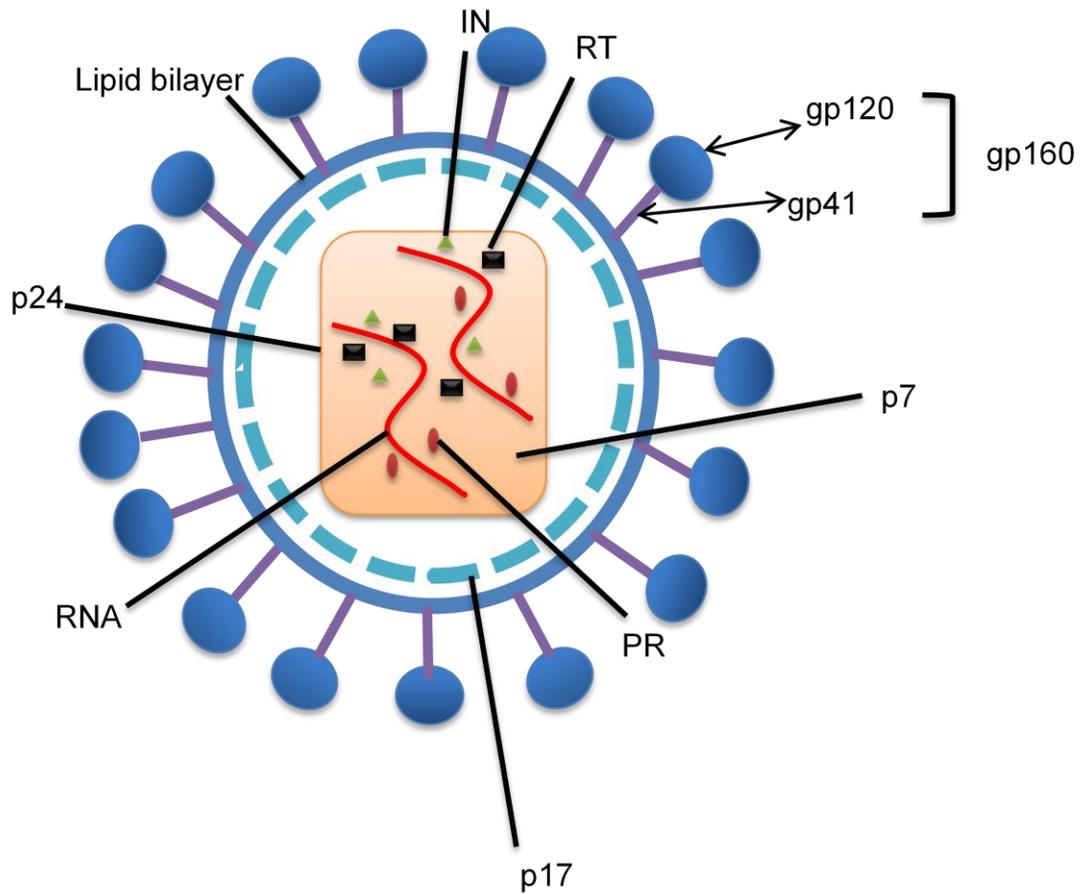


Figure 3: Schematic of HIV-1 structure, highlighting the important structural features associated with the virus: IN (integrase), RT (reverse transcriptase), PR (protease), RNA (genetic material), p24 (capsid protein), p7 (nucleocapsid protein), p17 (matrix), gp120 and gp41 (envelope proteins) and lipid bilayer. Figure is not drawn to scale.

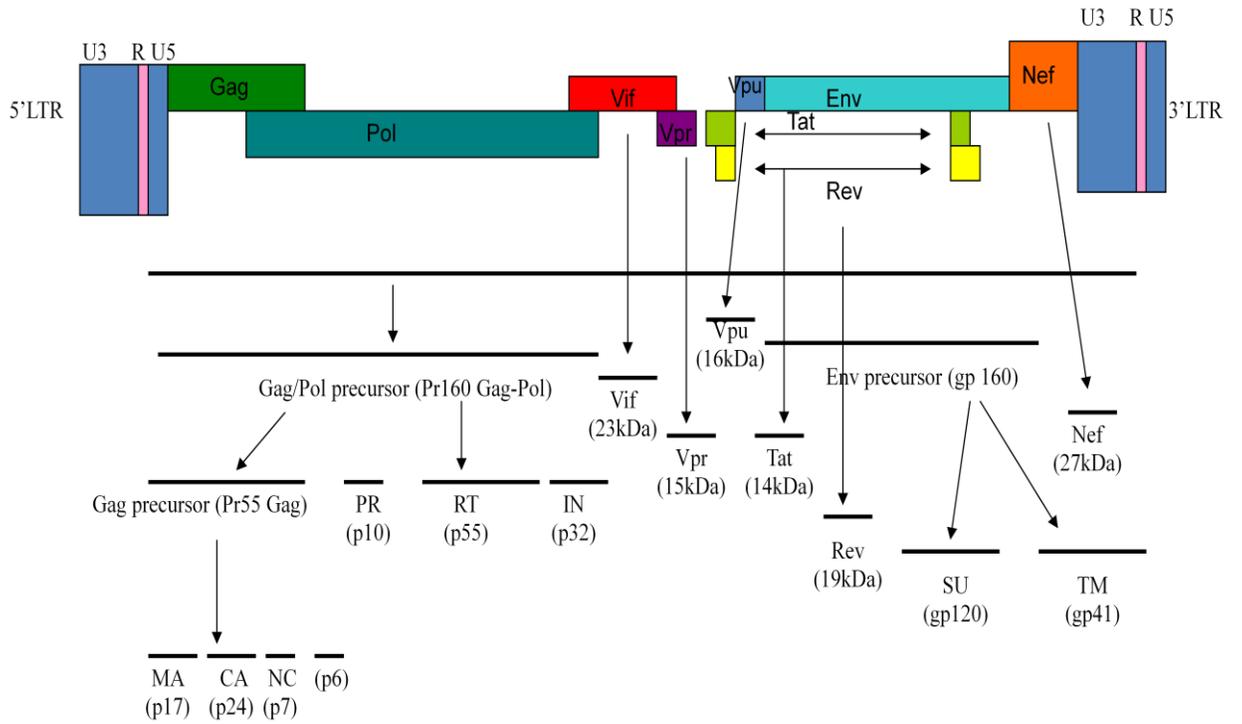


Figure 4: Genomic organization of HIV-1 and the processes of transcription and translation of HIV-1 genes. Figure is not drawn to scale.

after the 5' methylguanosine cap and at the 3' end of the RNA before the 3' Poly (A) tail. Upstream of the R region lies the unique 3' sequence (U3) which binds the majority of transcription factors such as nuclear factor kappa B and plays a central role in the activity of the LTR and viral expression (113). Downstream of the R region is another sequence known as the unique 5' sequence (U5), which includes one of the attachment (att) sites required for proviral integration (120). The order of the U3, R, and U5 sequences in the HIV-1 genome comprises the LTR (113).

Following the U5 sequence is the primer binding site (pbs), an eighteen nucleotide sequence, at which a host transfer RNA (tRNA) is hybridized to the genome and this is the site of initiation of minus-strand DNA synthesis (discussed in section 1.3.10). The region downstream from the pbs often contains the major signals for the encapsidation of the viral RNA into the virion particle in sequences called the packaging element (psi). The region also often contains a major splice donor site for the formation of subgenomic mRNAs. There are two short purine rich sequences located near the 3' end of the RNA genome, the polypurine tracts (PPT), which play a vital role in reverse transcription (189). These purine sequences, cPPT and 3' PPT, are located within the integrase gene and the upstream of the 3' LTR, respectively (189). The bulk of the RNA sequences that lies between the pbs and 3' PPT are coding regions for the viral proteins. Figure 5 illustrates these important sequences.

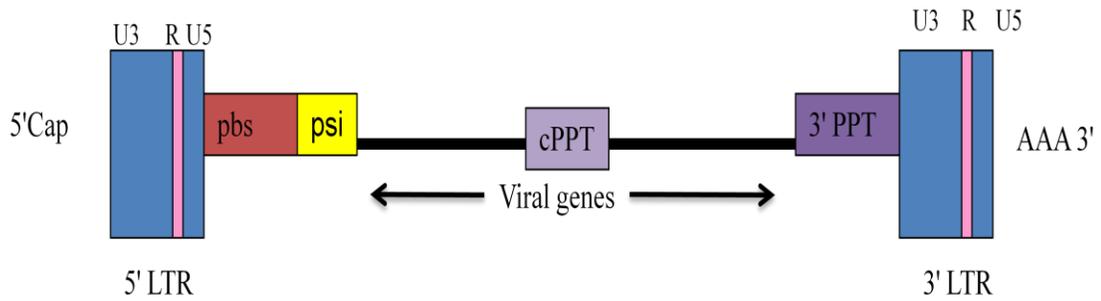


Figure 5: Important sequences required for HIV-1 replication. Each RNA molecule consists of a 5' cap at the beginning of the RNA molecule and has a sequence of poly (A) tails located at the 3' end of the RNA molecule. Each long terminal repeat (LTR) flanking the 5' and 3' viral genome is comprised of U3, R, and U5 sequences required for replication and integration into a host's genome. The packaging element (psi) is required for viral encapsidation. The primer binding site (pbs), central polypurine tract (cPPT), and 3' polypurine tracts (3' PPT) are sites important for the initiation of minus-strand DNA synthesis and plus-strand DNA synthesis, respectively.

1.3.4 HIV-1 Gag

The HIV-1 Gag proteins are necessary for a fully functional infectious virion (56, 72). The *gag* precursor encodes four Gag proteins (p17, p24, p7, and p6) which have numerous roles in the viral lifecycle (involved in virion assembly and maturation) and it is involved in the early post-entry steps in virus replication (56). The first protein, p17, is known as the matrix protein is involved in several stages during both the early and late stages of the viral replication cycle (52) as well as playing a structural role in anchoring the envelope proteins of HIV-1. The second capsid protein, p24, is involved in the formation of the capsid for HIV-1. Thirdly the nucleocapsid protein, p9, functions to bind to HIV-1 RNA and holds it in place in the virus core. Lastly, p6 assists the viral accessory protein, Vpu, in incorporation of newly made virions.

1.3.5 HIV-1 Pol

The *pol* gene is the most conserved region of HIV-1 that encodes three viral enzymes: protease (PR), reverse transcriptase (RT) (embedded with ribonuclease H (RNaseH)), and integrase (IN) (89, 174). PR is involved in proteolytic processing of the viral genome and cleaves the precursor viral polyprotein into smaller mature viral proteins. The viral enzyme, RT, is a RNA-dependent DNA polymerase that plays an important role in replicating the viral genome by reverse transcribing the viral ss RNA to ds proviral DNA (41). The process that RT plays in transcribing the viral RNA into DNA will be discussed in a later section. The function of the 32 kDa IN protein is to incorporate or "integrate" the proviral DNA into the host cell genetic material (57). During integration several nucleotides are removed from both 3' ends of the viral linear

DNA to generate a ds viral DNA molecule with 3' recessed ends (57). Staggered cuts are made in the host DNA by IN where the 3' recessed ends of viral ds DNA are joined together known as strand transfer (57). The process is completed when the host cellular enzymes repair the gaps between the host and viral genomes.

1.3.6 HIV-1 Env

The *env* gene encodes the viral envelope glycoproteins, gp120 and gp41, which are cleaved from the precursor protein gp160 (7, 24, 174). The gp41 protein is a membrane spanning protein that plays an important role in initiating the fusion between the host and cellular membranes, while the gp120 protein acts as a receptor binding molecule that is anchored by gp41 (refer to Figure 3) (50, 207). Gp120 and gp41 interact together non-covalently to form the HIV-1 envelope trimeric spike (each made up of three molecules of gp120 and gp41) used for entry into target cells. Both proteins are the targets of humoral immune response, especially gp120, and therefore gp120 is the most variable HIV protein (22, 156). The viral envelope proteins are essential for HIV-1 entry into host cells and play an important role in the type of cell HIV infects (called HIV tropism) based on the host cell receptor (50, 80, 82, 207).

1.3.7 HIV-1 accessory/regulatory proteins

The HIV-1 accessory proteins (Vpu, Nef, Vpr, and Vif) are known to play an important role in HIV-1 pathogenesis (6, 67) and each protein has a multifunctional role in HIV-1 infection (52, 53, 137). Vpu encodes a small transmembrane protein that enhances the release of infectious progeny virions from infected cells and includes the

downregulation of MHC I and CD4 receptors (47, 194). The most well known function of Nef is the downregulation of CD4 receptors (67) via binding of the cytoplasmic tail (53, 67, 159). This process is important as it allows the release of new virus particles (115). Other functions of Nef (not well understood) include class I MHCI downregulation, enhanced viral infectivity, and p21 activated protein kinase (PAK2) activation (53, 67, 159). Functions mediated by Vpr include LTR transactivation, nuclear import of the preintegration complex, and regulating apoptosis (6). The induction of cell cycle (G2 phase) arrest is mediated by the function of Vpr (6, 46). The function and role of HIV-1 Vif protein will be discussed in greater detail in Section 1.3.8 and subsequent chapters of this thesis.

The regulatory proteins, Tat and Rev, are viral regulatory factors required for HIV-1 gene expression (187). Rev plays an important role by binding to the Rev Responsive Element (RRE) located in HIV-1 *env*. It is responsible for producing viral structural protein by exporting genomic, unspliced, and partially spliced RNA from the nucleus (35, 176, 179). Tat is a small RNA protein responsible for HIV-1 transcription (160, 180) and like the accessory proteins, it is a multifunctional protein (164). Other roles of Tat include: modulating upregulation and downregulation expression of cellular genes (cytokines), ability to upregulate non-HIV viruses, immune suppression, and apoptosis (164).

1.3.8 HIV-1 Vif

HIV-1 Vif is encoded by all lentiviruses except for equine infectious anemia virus (20). It was first characterized in 1986 as a small one hundred ninety-two residue

cytoplasmic protein, with a molecular weight of 23 kDa (17, 20, 133, 188, 215). The C-terminal end of Vif consists of highly conserved cysteine residues located in three regions (position 114, 133, and the S¹⁴⁴LQXLA motif) (17) which are critical for viral infectivity and replication *in vivo* (61, 70, 188, 215). One critical function of HIV-1 Vif protein, *in vivo*, is that it increases the virus infectivity during the late stages of infection by ten-to-one thousand fold (61, 70, 188). Early studies showed that mutant HIV-1 virions lacking the *vif* gene or having deleterious mutations in the *vif* gene (Δ Vif) were approximately one thousand times less infectious than wild-type (wt) virions (61, 92, 188). It is known that Δ Vif could only replicate in certain cell types known as permissive cells (293 T, HeLa, SupT1, Jurkat) (20, 215). These cells support the replication of Vif-defective HIV-1 virions but several other non-permissive cell types have been identified that cannot support the replication of Vif-defective HIV-1 (certain CD4⁺ T cell lines, peripheral blood lymphocytes, macrophages, and H9 T cells) (17, 20, 215).

Over the last few years, HIV-1 researchers have been focussing attention on the Vif protein due to one of the most important function it serves - the ability to evade powerful antiviral activities of innate host molecules, **apolipoprotein B** mRNA editing enzyme, **catalytic polypeptide-like 3** (APOBEC3). The Vif protein has sequences located at the amino terminal and the carboxyl terminal domain that play important roles in inhibiting APOBEC3 degradation (17). The Vif protein induces degradation of APOBEC3 proteins by trafficking APOBEC3 proteins into the cellular ubiquitin-proteasome pathway. Vif acts as a physiological barrier where it prevents APOBEC3 molecules from being incorporated in newly budding virions (78, 178, 217, 218). The mechanisms and interactions of Vif and APOBEC3 proteins will be discussed in detail in

a later section of this thesis.

1.3.9 HIV-1 entry and reverse transcription

HIV-1 replicates through an extraordinary, complex, and unique lifecycle differentiating it from non-retroviruses. Each virion has two copies of genomic RNA, but upon entry into the host cell, this RNA is reverse transcribed into DNA that is then integrated into the host chromosomal DNA. The integrated form of the viral DNA, a provirus, serves as a template for the formation of viral RNAs and proteins that assemble progeny virions. These features of the lifecycle, especially the reverse flow of genetic information from RNA to DNA, and the integration of the DNA are the defining hallmarks of retroviruses.

HIV-1 entry into host cells occurs by first initiating receptor binding and membrane fusion to the plasma membrane of target cells. Binding occurs by envelope glycoproteins, gp120, present on the surface of the virus that binds to specific cells that have CD4 receptors on the cells surface. Binding of gp120/CD4 results in conformational changes to gp120 that allows gp120 to bind to chemokine co-receptors, CCR5 or CXCR4 (34, 207). Interactions with gp120 and chemokine receptors alter the transmembrane glycoprotein of HIV-1, gp41, which initiates the process of membrane fusion between the virus and host cell (124). This results in internalization of the viral core into the host cytoplasm of the infected cell where reverse transcription begins. This is a complex process that involves transcription of the RNA genome to ds linear DNA. This takes place in a large complex known as the reverse transcription complex (RTC) that roughly resembles the viral core and contains the following Gag and Pol proteins: Matrix (MA),

Nucleoprotein (NC), RT, and IN (16, 57). Following the release of viral nucleic acids (consisting of capsid and core proteins) from the viral core, a process of uncoating occurs which is still not completely understood (16, 57). Next, viral linear DNA is integrated to form the provirus. Then, transcription of the provirus to form viral RNAs is followed by splicing and nuclear export of the RNA. RNA is then translated to form precursor proteins which are assembled to the virion and packaging of the viral RNA genome, budding and release of the virions, and proteolytic processing of the precursors and maturation of the virions, where the virus uses host machinery to make newly virion particles.

After the formation of ds DNA, the proviral genome must enter the nucleus for incorporation into the host genome and entry occurs through a preintegration complex (PIC) (57, 216). The PIC involves a nucleoprotein complex made entirely of a mixture of viral DNA, host, and viral proteins (Vpr, MA, RT, IN) (8). Once in the nucleus, genomic replication can begin.

1.3.10 HIV-1 genome replication

The process of HIV-1 replication involves a combination of host cellular and viral molecules resulting in a ds viral genome capable of permanently integrating into the host genome. Each step of replication will be discussed below (Figure 6).

The process of reverse transcription is initiated from a paired 3' OH host cellular primer tRNA, tRNA^{Lys}, that anneals to a complementary sequence on the viral RNA genome, the pbs (110, 189). DNA is first synthesized from the pbs primer, using the plus-strand RNA genome as a template, and moves toward the 5' end of the RNA to form

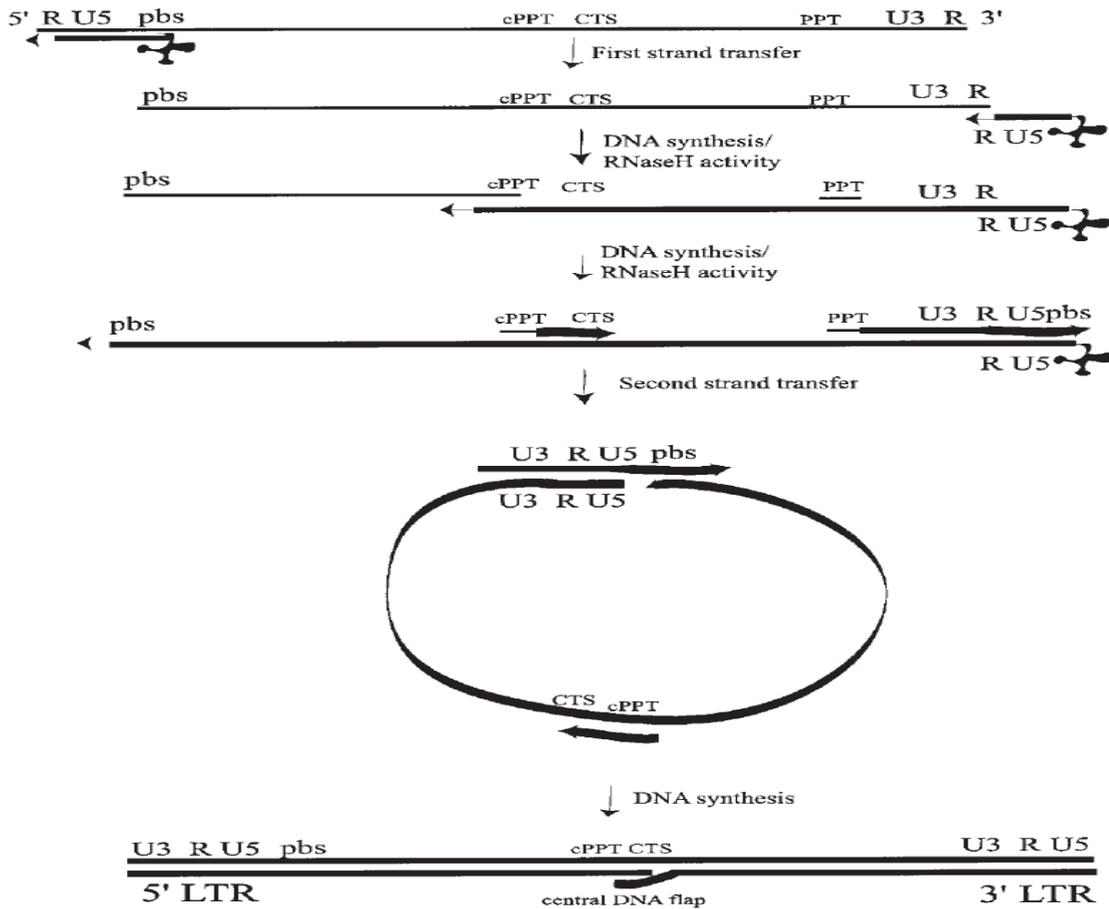


Figure 6: Highlights of HIV-1 reverse transcription in the process of strand synthesis (57). Several steps are involved in reverse transcription of the HIV-1 genome. The important features highlighted are the primer binding site (pbs) which is required for initiating reverse transcription in the 5' to 3' direction, RNaseH activity is required to degrade RNA from the RNA:DNA hybrid, purine rich sequences resistant to RNaseH (cPPT and 3' PPT). Thin black lines are indicated as RNA whereas thick lines are DNA. Figure reproduced with permission from Dr. Eric O. Freed from the National Drug Resistance Program and National Cancer Institute in Frederick, MD. Permission obtained January 17, 2010.

minus-strand DNA sequences (90, 189). A RNA: DNA hybrid is formed where the primer remains attached to its 5' end and most of the RNA is degraded by the RNase H activity of RT. This exposes the ss DNA before it enters the plus-strand DNA synthesis (23, 57).

Two short purine rich sequences near the 3' end of the genome, the cPPT and 3' PPT, are relatively resistant to the activity of RNase H and serve as the primers for synthesis of the plus-strand DNA (23). The oligonucleotides that remain hybridized to the minus-strand DNA serve as the primers for synthesis of the plus-strand DNA, using the minus strand DNA as template (23, 189). Once the primers have served to initiate DNA synthesis, they are quickly removed from the DNA by RNase H (57). Synthesis proceeds toward the 5' end of the minus strand, first copying the U3, R, and U5 sequences, and then extending to copy a portion of the primer tRNA still present at its 5' end. Elongation stops and removal of primer tRNA at the 5' end of the minus-strand DNA occurs by RNase H.

The removal of the tRNA exposes the 3' end of the plus strand DNA to pair with the 3' end of the minus-strand DNA. The sequences anneal via the shared pbs sequences and forms a circular intermediate where both strands are elongated. The final extension of the minus-strand DNA is coupled to displacement of the plus-strand strong-stop DNA from the 5' end of the minus strand; as minus-strand elongation occurs; the plus strand strong stop is peeled away and transferred to the 3' end of the minus strand. At the end of this elongation, the circle is opened up into a linear DNA. The plus strands are all extended and displacement synthesis may occur to remove short DNAs and make longer plus-strand DNAs. Essentially, now the genome is a ds DNA molecule where it integrates

into the host genome and this is called a proviral intermediate.

1.3.11 Viral diversity in HIV-1

HIV-1 constantly develops viral mutations due to an error-prone mechanism induced when the virus replicates its viral genome. This results in high mutation rates (39, 119, 129). This cloud of viral diversity results in HIV-1 viral quasispecies, both within an individual but also at the population level (119). The replication error frequency within a host is 3.4×10^{-5} mutation sites⁻¹ generation⁻¹ and the production rate of the virus *in vivo* exceeds 10^{10} virions per day (122). This results in large numbers of unproductive virus, but also contributes to rapid evolution in viral fitness, immune evasion, and drug resistance. This of course makes vaccine and HIV drug development very difficult to achieve.

1.4 HIV-1 host restriction factors

Several innate host factors have recently been identified that contribute to reduce HIV-1 disease progression and these host factors play an important role in fighting HIV-1. Some such as bone marrow stromal antigen 2 (BST-2) and tripartite motif 5-alpha (Trim5- α) are potent inhibitors that are constitutively expressed on B cells, dendritic cells, and activated T cells (114, 137, 194). BST-2 is involved in counteracting the function of HIV-1 Vpu by inhibiting mature virion release of HIV-1 through tethering the virions to the host cell surface (47, 114, 137, 194). As for Trim5- α , it is a restriction factor that localizes in clusters in the cellular cytoplasm (18). It plays an important role in species-specificity by restricting HIV-1 by interfering with the post-entry steps involved in capsid

disassembly (18, 69, 79, 102, 137). Over several years, considerable attention has been made to a special group of host restriction factors known as the human family of APOBEC cytidine deaminases (20, 64, 189).

1.4.1 AID / APOBEC family

Members of the activation-induced deaminase (AID)/APOBEC family are a group of cytoplasmic proteins which exhibit diverse physiological functions and act by deaminating cytidine (C) to uracil (U) in RNA and/or DNA (51, 93, 151, 189, 195). The AID/APOBEC family consists of: AID, APOBEC1, APOBEC2, APOBEC3, and APOBEC4 subgroups, which are encoded on different chromosomes in humans. AID and APOBEC1 are located in close proximity to each other and are located on chromosome twelve, while the others are encoded on different chromosomes. Each AID/APOBEC family member has different specific functions. For instance, AID is involved in antibody diversification (somatic hypermutation, gene conversion, and class-switch recombination) by deaminating cytosine residues in the DNA of the immunoglobulin locus while APOBEC1, the first member to be discovered in this family, is highly expressed in the tissues of the small intestine (25, 32, 33, 78, 142) and is responsible for editing the mRNA for a lipid transport protein, apolipoprotein B. APOBEC1 induces deamination at position C6666 (coding for glutamine) to a stop codon, essentially creating a shorter version of apolipoprotein B protein (25, 32, 204). This editing allows the transportation of triglycerides from the small intestine to the tissues (25, 32). APOBEC2, located on chromosome six, is highly expressed in muscle tissue (32, 142) and has an unknown function (51). APOBEC3 is composed of seven related genes found in a cluster on

chromosome twenty-two and it was thought to result from gene duplication (17, 142). They include: APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3DE, APOBEC3F, APOBEC3G, and APOBEC3H. Members of the APOBEC3 cluster encode proteins that are structurally and functionally related to the C to U RNA-editing cytidine deaminase APOBEC1 (133). It has been well documented that the APOBEC3 proteins have potent anti-viral properties against a wide range of human and animal viruses such as: HIV, hepatitis B virus, human T-cell leukemia virus type 1 (HTLV-1), mouse mammary tumor virus, adeno-associated virus, and various endogenous retroviruses (12, 32, 51, 85, 127, 185). APOBEC3 proteins are also known to inhibit a variety of endogenous retroelements (32, 51, 142). APOBEC4, primarily found in the testes, is located on chromosome one (142) and like APOBEC2 has an unknown function (32).

All members in the APOBEC family share a conserved cytidine deamination motif, His-x-Glu-x₂₃₋₂₈-Pro-Cys-x₂₋₄Cys, (where “x” denotes any amino acid) and either have one or two zinc binding domains within the active site that is involved in cytidine deamination (17, 85, 93, 142, 146, 157, 189). The mechanism of APOBEC-dependent cytosine deamination occurs when a zinc ion interacts with the histidine and two cysteine residues along with a glutamic acid residue (32, 78, 85). A sequence of protonations and bond cleavages between the active site and cytosine occur when a water molecule reacts with a zinc ion and glutamate producing a hydroxide ion resulting in the release of the amino group from the cytosine (78, 85). APOBEC1, AID, APOBEC2, APOBEC3A, APOBEC3C, APOBEC3H, APOBEC4 all contain a single deaminase domain whereas APOBEC3B, APOBEC3DE, APOBEC3F, and APOBEC3G contain two deaminase domains in their active sites (17, 78).

1.4.2 APOBEC3G/F in HIV-1 infection

The two best studied APOBEC3 proteins, APOBEC3G and APOBEC3F, are present in cells non-permissive for HIV-1 infection and are involved in counteracting the increased replicative ability caused by HIV-1 Vif (133). APOBEC3G and APOBEC3F have a 50% amino acid sequence identity (12). APOBEC3G demonstrates the strongest anti-viral activity in HIV-1, whereas APOBEC3F has a weaker anti-viral activity (75, 203, 212). During packaging of new virion particles (in the absence of Vif), these cellular defence molecules are "hijacked" into budding viruses by binding to Gag proteins and viral RNAs (2, 17, 19, 127) and carry out their deamination functions when the newly made virus infects a new target cell (Figure 7). In the absence of Vif, APOBEC3G binds to the nucleocapsid NC protein of Gag and incorporates itself into the virion in close proximity to the reverse transcription complex (70). APOBEC3G and APOBEC3F catalyze cytosine deamination during the process of minus-strand DNA synthesis (133). Essentially, one of two processes occurs where either viral DNA is degraded or lethal proviral hypermutation occurs (not always) due to altering the amino acid sequences at the HIV-1 proviral genome (12, 133, 196). The C-terminal domain of these APOBEC proteins targets C to U substitutions in the minus strand DNA of HIV-1. This creates an inordinate number of G to A transitions resulting in several stop codons in the reverse transcript, often leading to the production of incompetent viruses and at times abolishing production of viral progeny (78, 118, 165).

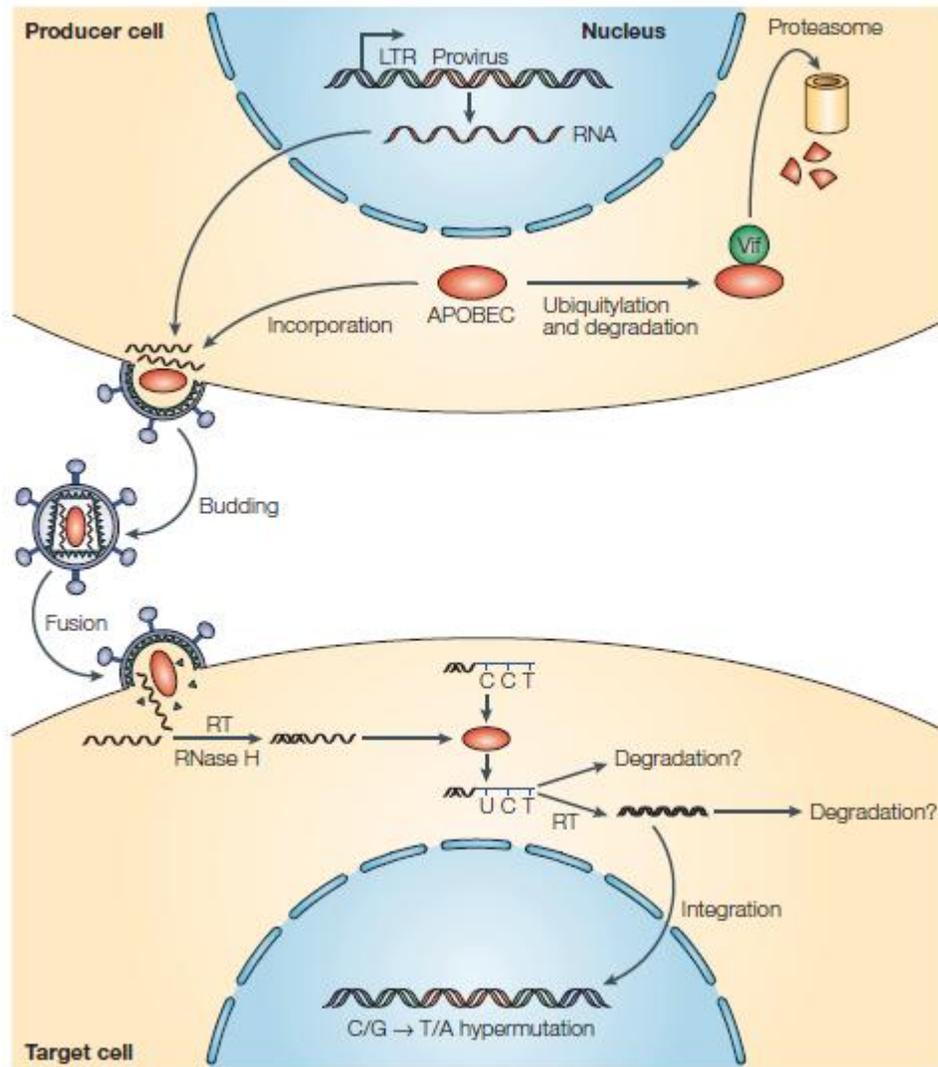


Figure 7: Molecular mechanisms of APOBEC proteins and Vif can influence the retroviral lifecycle (78). In the producer cell, APOBEC proteins can be ubiquitinated and degraded in the proteasome by HIV-1 Vif or APOBEC proteins can be packaged into new virions. APOBEC proteins that are successfully packaged into new virions can induce G to A proviral hypermutation in the target cell. Figure reproduced with permission from Dr. Reuben Harris. Permission obtained January 9, 2011.

Other mechanisms of anti-HIV activity by APOBEC3G/F proteins were first discovered by conducting site-directed mutagenesis as well as deletion studies of the APOBEC catalytic site. It was found that APOBEC proteins are able to undergo a process known as deaminase-independent restriction (84, 88, 92, 140, 173, 181, 186). This process was first described in HTLV-1 studies where HTLV-1 infection was inhibited by a method other than cytidine deaminase activity (173). Although some studies propose APOBEC inhibition of HIV-1 in an independent manner free from G to A hypermutation, this mechanism is unclear and controversial as some studies suggest that cytidine deaminase activity is a necessary function (130, 173). This deaminase-independent restriction is thought to involve APOBEC binding to viral RNA and hence, preventing the activity of RT to elongate the viral RNA to DNA (6, 206). Also, it has been shown that the dose dependent action of deaminase-independent restriction relies on the high expression of APOBEC3 proteins (6, 206).

1.4.3 Mechanism of HIV-1 Vif inhibition of APOBEC3G/F

Vif recruits the host E3 ubiquitin ligase complex which is composed of cellular proteins (elongin B, elongin C, cullin 5, and RING-box 2) that are essential for the induction of APOBEC3G and APOBEC3F polyubiquitination and proteasomal degradation (11, 127, 133, 178) (Figure 8). The process by which Vif binds the E3 ubiquitin ligase complex involves three critical sequences at the C terminal region of the Vif protein. The BC box of vif (positions 144 to 150) directly binds to elongin C, while the zinc-binding region at positions 120, 123, and 124 of Vif bind to cullin 5 through hydrophobic bonds via a highly conserved motif, H¹⁰⁸_{X₅}CX₁₇₋₁₈CX₃₋₅H¹³⁹ (17, 217). The

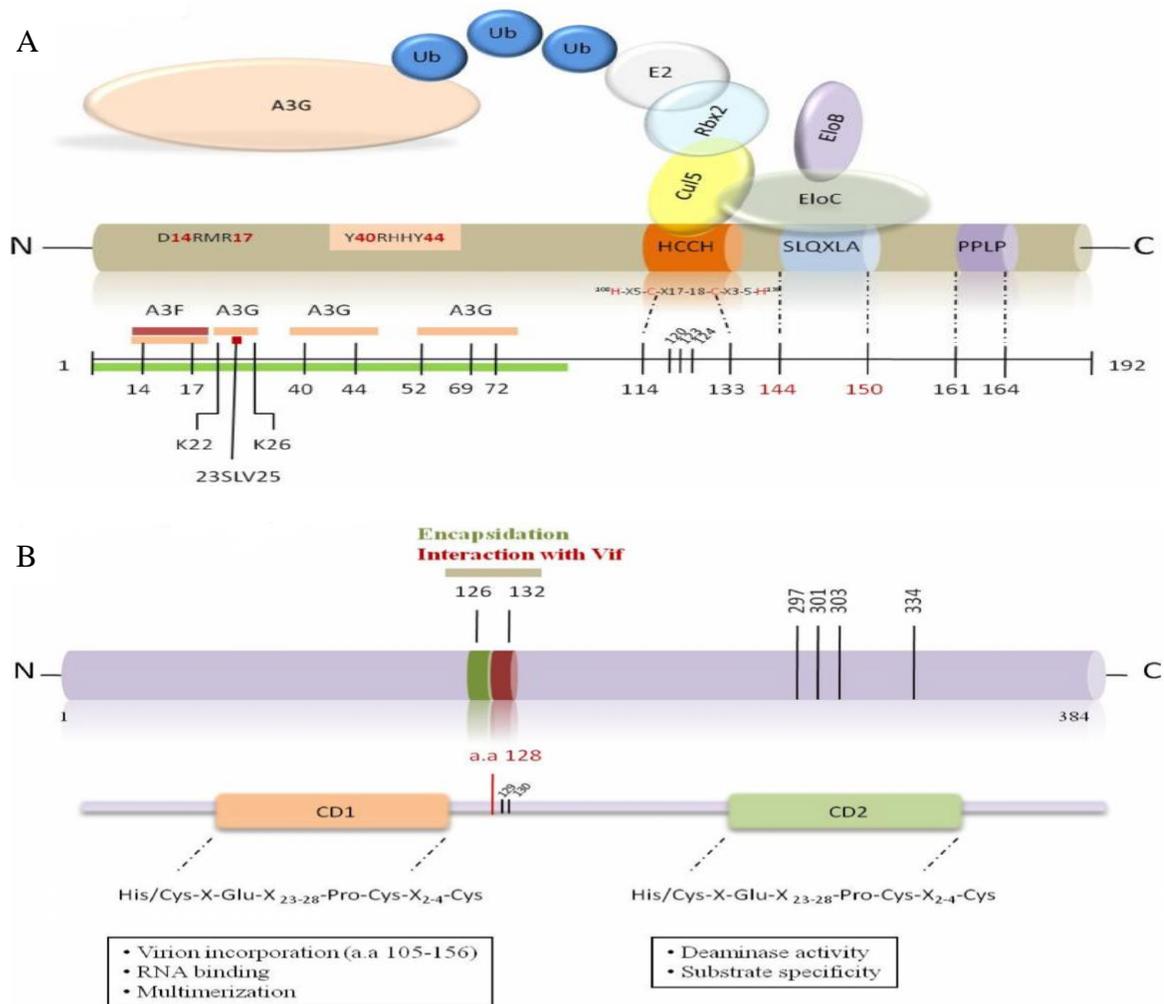


Figure 8: The protein interactions between Vif and APOBEC3G (A3G) domains (17). (A) HIV-1 Vif domains are highlighted in pink, orange, and light blue where each represents specific amino acid residues (N-terminal region), zinc binding HCCH domain, and SLQXLA domain, respectively. The N-terminal domain of Vif is involved in A3G interaction while the C-terminal region targets A3G proteasomal degradation. (B) The A3G cytidine deaminase activity is found in the CD1 and CD2 catalytic domains. Amino acid position 126 in A3G is responsible for its encapsidation and position 132 is a region to where Vif binding occurs. Figure reproduced with permission from Dr. Joao Goncalves. Permission obtained February 4, 2011.

H¹⁰⁸X₅CX₁₇₋₁₈CX₃₋₅H¹³⁹, along with a second sequence, S¹⁴⁴LQxLA¹⁴⁹, are responsible for APOBEC degradation (17, 196, 217). Another sequence in Vif, the Y⁴⁰RHHY⁴⁴, is essential for the interaction with APOBEC3G (196). On average, there are about 7 ± 4 APOBEC molecules per virion is packaged in Δ Vif (208). The binding of Vif to APOBEC3G/F decreases the intracellular concentration of APOBEC3G and APOBEC3F, preventing their incorporation into newly formed HIV-1 virions (133, 218).

1.4.4 APOBEC3G/F-mediated hypermutation of the HIV-1 genome

The mechanism by which APOBEC3G/F interacts with the viral genome first begins with it altering the viral minus strand DNA by creating excessive levels of U, through an APOBEC-dependent deamination of C residues (12). Then, adenine (A) is incorporated into the plus DNA strand due to RT recognizing thymine (T) from the alteration to U on the minus strand DNA (12, 78). This results in several guanine (G) to A nucleotide point mutations in the plus strand of the HIV-1 genome (77, 121, 128, 214).

A study conducted by Harris and colleagues identified that proviral DNA in the presence of APOBEC3G showed extensive G to A hypermutation along proviral DNA, where the level of guanines consisted of 25% of the total nucleotides present compared to proviral DNA sequences that did not have APOBEC3G present (77). The changes that occur in the minus strand sequence as a result of APOBEC3G and APOBEC3F activity differ slightly in terms of initial deamination because each protein has different sequence specificities. Cytidine deamination for APOBEC3G occurs at a distinct dinucleotide sequence, CC, while APOBEC3F prefers to target TC nucleotide bases (78, 189). The resulting minus strand DNA synthesis for APOBEC3G is CC to CT and APOBEC3F is

TC to TT (78, 189). Due to APOBEC3G and APOBEC3F sequence specificities, the plus strand DNA sequences have a pattern of GG to AG and GA to AA substitutions, respectively (78, 146, 165, 189, 196).

In 2006, Suspene *et al.* explored the distribution of APOBEC3G/F editing across the HIV-1 genome of thirty sequences. They formulated an equation where they were able to find two highly polarized gradients where APOBEC-mediated hypermutation was proposed to occur on the HIV-1 genome (Figure 9). The two gradients, one near the 5' cPPT and 3' LTR, are significant regions during HIV-1 replication as they correspond to the duration the minus strand remains ss during replication (189). The two high APOBEC3-mediated hypermutation gradients at the HIV-1 genome suggested that APOBEC3 molecules varied significantly upon packaging (189). Thus, the pol region may show the most affects of APOBEC-mediated hypermutation (189).

1.5 The significance of CD4⁺ T cell and viral load counts in HIV-1 disease progression

There are two important measurements for monitoring HIV-1 disease progression (161, 205). The first measurement, CD4⁺ T cell count, measures the levels of CD4⁺ T cells/mL of blood. CD4⁺ T cells are the main target of HIV-1 and more importantly the key cells that are required to fight off pathogens (68). CD4⁺ T cell counts can be a good predictor in determining how far along a HIV-1 infected individual is towards developing AIDS (83). Most importantly, CD4⁺ T cell counts are commonly used as a marker of disease progression in cases where viral load (discussed below) is not measured.

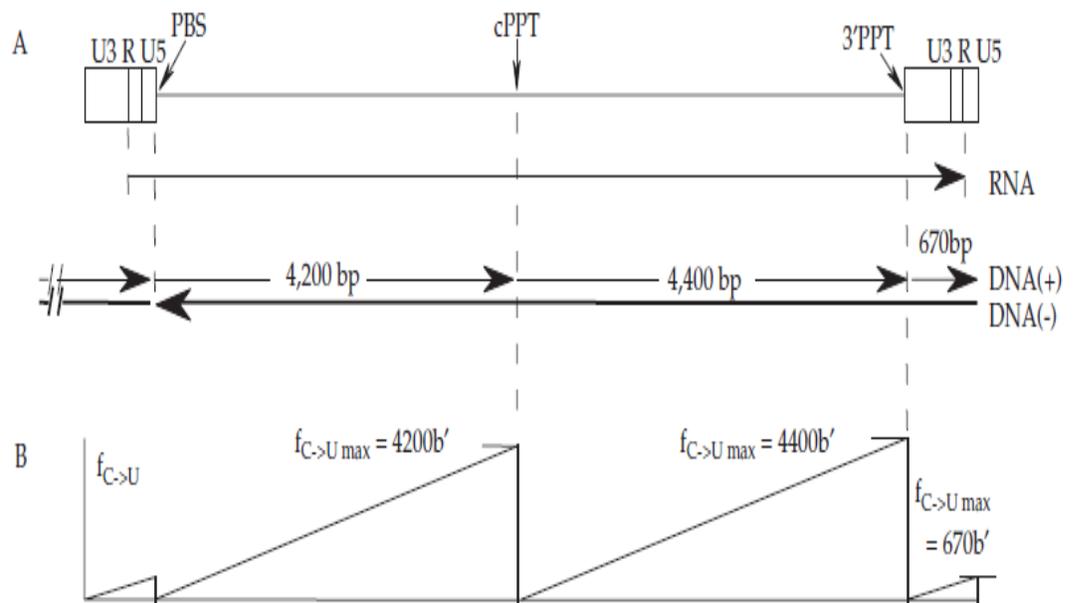


Figure 9: Two distinctive twin gradients are observed during editing of the HIV-1 DNA genome by activity of APOBEC3 proteins (189). (A) Drawn to scale is the first step of transcription of the HIV-1 genome by reverse transcription. The important features required for reverse transcription are: pbs (primer binding site), cPPT (central polypurine tract), and 3' PPT (3' polypurine tracts). (B) Two distinctive twin gradients are formed in the HIV-1 genome, illustrating the different rates and locations where cytosine deamination would occur due to APOBEC-mediated hypermutation in the HIV-1 genome, particularly more frequent in the pol region. Figure reproduced with permission from Dr. Simon Wain-Hobson. Permission obtained February 2, 2011.

The second measurement, plasma viral load, is a quantitative measure of viral nucleic acid in the blood of a HIV-1 infected individual and is the single best predictor for monitoring HIV-1 disease outcomes (68, 123, 145). Viral load measures the number of RNA copies/mL of blood. A high viral load is due to the virus quickly replicating and therefore the rate of HIV-1 disease progression may be accelerated (144). A high viral load lies within the range of 5000 - 1×10^6 RNA copies/mL of blood and is associated with progression to AIDS (144). Individuals with high viral loads who are progressing to AIDS may be placed on ARV therapy to control the levels of virus in the blood (197).

Viral load data is a good indication in determining the disease progression rates of individuals with HIV-1 infections (123), but in some cases viral loads are at undetectable levels. This is due to the viral RNA is not reaching the minimum threshold for detection. This does not mean the infection is cleared; the viral genetic material is latently integrated into the host cell genome (161). The incorporation of viral nucleic acid into the host genome during a persistent infection can affect viral load as it does not reflect the amount of integrated virus and therefore; may have a different effect on predicting disease progression and treatment outcome depending on the stage of the disease (Rozeza 2010). Furthermore, the detection of proviral DNA is critical in following HIV-1 disease progression (161, 205) as the amount of integration roughly may reflect viral load.

Ideally, monitoring the combination of both plasma viral load data and the number of CD4⁺ T cell count in an infected individual will provide information on how to monitor HIV disease status, give guidance to therapy options, and the long-term course of HIV-1 infection (83, 87).

1.6 APOBEC and disease progression

Heterogeneity of HIV-1 disease progression can be observed in any population and has been described in several sex worker cohorts (5, 96, 105). The rates of disease progression and pathogenesis within these cohorts vary substantially (5). The role of APOBEC3 proteins and correlation to disease progression are currently under debate in the scientific literature (5, 26, 97, 118). Without doubt, APOBEC3 proteins act as potent anti-viral host proteins that are able to cause excessive G to A hypermutation at various regions of the HIV-1 genome. Previous sequence analysis of a subset of HIV-1 infected, antiretroviral naïve women from the Pumwani Commercial Sex Worker (CSW) cohort from Nairobi, Kenya (54) found dramatic hypermutation in the vpu/env region of proviral HIV-1 DNA ($p \leq 0.05$) (118). Furthermore, women from the CSW cohort with hypermutation at vpu/env region had a significantly higher CD4⁺ T cell count than women of the same CSW cohort without hypermutation (118).

2.0 Rationale, hypothesis and objectives

Rationale

The link between APOBEC3G/F activity and disease progression are still unclear even given the data from Land *et al.* that shows that an increase in hypermutation in HIV-1 Vif in patients with higher CD4⁺ T cell counts. Based upon Figure 9 and the work of Suspene *et al.*, I expect to also observe an increase in hypermutation in the HIV-1 pol region (189). Thus, I propose that individuals with slower disease progression as measured by CD4⁺ T cell count will have an increased frequency of APOBEC-mediated hypermutation in the pol region of HIV-1. As well I will also determine HIV-1 Vif activity from viruses in subjects that exhibit hypermutation to determine if they have impaired expression and ability to inhibit APOBEC3G. Besides the study by Land *et al.*, there is a lack of information on the link between CD4⁺ T cell counts and hypermutation in other regions of HIV-1 provirus. Here, I will build on previous findings to further characterize and identify APOBEC3G-mediated hypermutation in HIV-1 proviral pol region and its relationship with CD4⁺ T cell counts. In doing so, I hope to gain some insight in why some individuals progress faster or slower to AIDS and whether the function of APOBEC3G-mediated hypermutation plays a role. Such knowledge will contribute to the development of effective prophylactic and therapeutic modalities for HIV-1 infection.

Based on the above information, we hypothesize that:

Individuals with slower HIV-1 disease progression as measured by elevated CD4⁺ T cell counts will show an increase in hypermutation in HIV-1 pol; as well as in the vpu/env region as previously described. This will be due to infection by viruses containing HIV-1 Vif that is deficient in its ability to affect APOBEC3G function.

To investigate this hypothesis, four specific objectives will be completed:

Objective 1: Sequence a segment of the *pol* gene in antiretroviral naïve patients in the Malaya (ML) cohort from Pumwani, in Nairobi, Kenya to analyze for the presence of APOBEC-mediated HIV-1 proviral hypermutation.

Objective 2: Determine the relationship between the presence of APOBEC3G/F-mediated hypermutation in the pol region and correlate with CD4⁺ T cell counts.

Objective 3: Compare the relative expression of APOBEC3G from peripheral blood mononuclear cells (PBMCs) of subjects with or without APOBEC-mediated hypermutation, using Western blot.

Objective 4: To determine, *in vitro*, how Vif, isolated from patients with high level of hypermutation in the HIV-1 env/vpu region affect wt APOBEC3G expression, when co-expressed in 293T cells.

3.0 Materials and methods

3.1 General laboratory materials

3.1.1. Laboratory prepared solutions

1% agarose gel

1 gram (g) UltraPure™ agarose, catalogue number (Cat. No.) 16500-500 (Invitrogen: Burlington, Ontario, Canada)

Ethidium bromide (EtBr) (10 milligrams (mg)/mL) (Sigma-Aldrich: Oakville, Ontario, Canada)

to a final volume of 100 mL 1x tris-borate EDTA (TBE) buffer

Luria-bertani (LB) - ampicillin agar plates

37.0 g LB agar miller powder (contains 10g peptone from casein, 5 g yeast extract, 10 g sodium chloride, 12.0 g agar-agar), Cat. No. 1.10283.0500 (EMD Chemicals: Darmstadt, Germany)

Ampicillin, Cat. No. 10 835 242 001 (Roche Diagnostics: Mannheim, Germany)

100 mg/mL

to a final volume of 1litre (L) sterile-filtered double-distilled (ddH₂O), autoclaved for 15 minutes at 121 degrees Celsius (°C) (STERIS Amsco Century: SG-120 Scientific Gravity Sterilizer)

LB - ampicillin broth

25.0 g LB broth miller powder (contains 10g peptone from casein, 5 g yeast extract, 10 g sodium chloride), Cat. No. 1.10285.0500 (EMD Chemicals: Darmstadt, Germany)

Ampicillin, Cat. No. 10 835 242 001 (Roche Diagnostics: Mannheim, Germany)

100 mg/mL

to a final volume of 1L sterile-filtered ddH₂O, autoclaved for 15 minutes at 121°C
(STERIS Amsco Century: SG-120 Scientific Gravity Sterilizer)

Mild stripping buffer

15 g glycine, Cat. No. 161-0718 (Bio-Rad: Mississauga, Ontario, Canada)

1 g sodium dodecyl sulfate (SDS), 10 mL tween-20

to a final volume of 1 L with sterile-filtered ddH₂O

Adjust pH to 2.2

Phosphate buffered saline (PBS) (pH 7.4)

48.5 g PBS powder (contains 137.93 millimolar (mM) NaCl, 2.67 mM KCl, 8.1 mM Na₂HPO₄, 1.41 mM KH₂PO₄), Cat. No. 2.1600-044 (Invitrogen: Burlington, Ontario, Canada)

to a final volume of 1L with sterile-filtered ddH₂O

Adjust pH to 7.4

Radioimmunoprecipitation assay (RIPA) cell lysis buffer for protein isolation

500 microlitre (µl) 20% SDS, 1mL triton x100, 0.5 g sodium deoxydriate, 0.6057 g tris
pH 8.0

to a final volume of 100 mL sterile-filtered ddH₂O

1 protease inhibitor cocktail tablet per 10 mL lysis buffer, Cat. No. 11 836 153 001
(Roche Diagnostics: Mannheim, Germany)

SDS loading buffer (2x)

80 µl glycerol, 50 µl 1M Tris-Cl (pH 6.8), 160 µl 10% (w/v) SDS, 20 µl 0.1 % (w/v)
bromophenol blue, 40 µl β-mercaptoethanol

TBE buffer (pH 8.13-8.23)

54.0 g Tris-base, 27.5 g boric acid, 2.92 g EDTA

to a final volume of 1.0 L sterile-filtered ddH₂O

Tris EDTA (TE) buffer (pH 8.0)

100 µl 1 M Tris-HCl pH 8.0, Cat. No. VW1500-01 (VWR International: Mississauga, Ontario, Canada)

20 µl 0.5 M EDTA (pH 8.0), Cat. No. 15576-028 (Invitrogen: Burlington, Ontario, Canada)

to 99.8 mL sterile-filtered ddH₂O

Tris buffered saline (TBS) (10x)

87.7 g NaCl, 12.1 g Tris, 4 mL HCl

to a final volume of 1 L sterile-filtered ddH₂O

TBS (1x) + 0.1% Tween-20

100 mL 10x TBS

1 mL Tween-20, Cat. No. H5151 (Promega: Madison, Wisconsin, U.S.A.)

to a final volume of 1 L sterile-filtered ddH₂O

Freezing media (2x)

20% dimethylsulfoxide (DMSO), 40% fetal bovine serum (FBS), 40% culture media

6x polymerase chain reaction (PCR) gel loading buffer

0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol in H₂O

3.1.2 Reagents

Antibodies

Anti-G3PDH/GAPDH rabbit polyclonal antibody, Cat. No. 2275-PC-100 (Trevigen: Gaithersburg, Maryland, U.S.A.)

Anti-mouse IgG-HRP, peroxidase-linked species-specific whole antibody (from sheep), Cat. No. NA931 (GE Healthcare/Amersham Biosciences: Sweden)

APOBEC3G Antiserum (ApoC17), Cat. No. 10082 (NIH AIDS Research & Reference Reagent Program: Germantown, Maryland, U.S.A.)

beta-Actin mouse monoclonal antibody, Cat. No. mAbcam 8224 (Abcam: Cambridge, Massachusetts, U.S.A.)

Goat anti-rabbit IgG HRP, Cat. No. 111-035-144 (Cedarlane: Burlington, Ontario, Canada)

HIV-1 BH10 Vif Antiserum, Cat. No. 809 (NIH AIDS Research & Reference Reagent Program: Germantown, Maryland, U.S.A.)

HIV-1 HXB2 Vif Antiserum, Cat. No. 2221 (NIH AIDS Research & Reference Reagent Program: Germantown, Maryland, U.S.A.)

Cell culture

DMEM media with L-Glutamine, Cat. No. D5671 (Sigma-Aldrich: Oakville, Ontario, Canada)

10% Fetal calf serum (heat-inactivated at 56°C for 30 minutes), Cat. Nos. Gibco 12483 and A12617 (Invitrogen: Burlington, Ontario, Canada)

Fungazone (HyClone: Logan, Utah, U.S.A.)

2% Penicillin and streptomycin, Cat. No. Gibco 15240 (Invitrogen: Burlington, Ontario,

Canada)

Roswell park memorial institute medium (RPMI), Cat. No. 12-702F (Lonza: Switzerland)

TrypLE™, Cat. No. 12604-021 (Invitrogen: Burlington, Ontario, Canada)

Cell separation

Trypan Blue, Cat. No. T8154 (Sigma-Aldrich: Oakville, Ontario, Canada)

DNA extraction and purification from whole blood

EZ1^R Kit, Cat. No. 1023745 (Qiagen: Mississauga, Ontario, Canada)

DNA extraction from PBMCs

QIAamp DNA Mini Kit (Qiagen, Mississauga, Ontario) was used for isolating genomic (including proviral) DNA

Plasmids

pcDNA-HVif, Cat. No.10077 (NIH AIDS Research & Reference Reagent Program: Germantown, Maryland, U.S.A.)

pCMV-HA, Cat. No. PT3283-5 (Clontech Laboratories: Mountain View, California, U.S.A.)

Restriction enzyme digest

GelPilot Loading Dye (5x), Cat. No. 239901 (Invitrogen: Burlington, Ontario, Canada)

pst I restriction enzyme, Cat. No. 15215-023 (Invitrogen: Burlington, Ontario, Canada)

xba I restriction enzyme, Cat. No. 15226-012 (Invitrogen: Burlington, Ontario, Canada)

10x React[®]2 buffer for pst I and xba I (Invitrogen: Burlington, Ontario, Canada)

1 kb Plus Ladder, Cat. No. 10787-018 (Invitrogen: Burlington, Ontario, Canada)

Transfection

TransIT[®] - LT1 Transfection Reagent, Cat. No. 2305 (Mirus: Madison, Wisconsin,

U.S.A)

Transformation

S.O.C. medium, Cat. No. 15544-034 (Invitrogen: Burlington, Ontario, Canada)

One shot[®] MAX efficiency[®] DH5 α TM – T1^R competent cells, Cat. No. 12297-016
(Invitrogen: Burlington, Ontario, Canada)

3.1.3. Commercial kits

Bicinchoninic acid (BCA) assay

Bovine serum albumin (BSA) standard (2 mg/mL), Cat. No. 500-0206 (Bio-Rad:
Hercules, California, USA)

BCA protein assay kit, Cat. No. 71285-3 (Novagen; San Diego, California, USA)

Gel electrophoresis

4x NuPAGE[®] LDS sample buffer, Cat. No. NP0008 (Invitrogen: Burlington, Ontario,
Canada)

10X NuPAGE[®] reducing agent (500 mM dithiothreitol), Cat. No. NP0004 (Invitrogen:
Burlington, Ontario, Canada)

NuPAGE[®] Novex[®] 4-12% Bis-Tris Gel 1.0 mm x 10 well, Cat. No. NP0342BOX
(Invitrogen: Burlington, Ontario, Canada)

NuPAGE[®] MES SDS running buffer (20x) Cat. No. NP0002 (Invitrogen: Burlington,
Ontario, Canada)

NuPAGE[®] antioxidant, Cat. No. NP0005 (Invitrogen: Burlington, Ontario, Canada)

Novex[®] sharp pre-stained protein standard, Cat. No. LC5800 (Invitrogen: Burlington,
Ontario, Canada)

Plasmid DNA purification

QIAprep spin miniprep kit (250) Cat. No. 27106 (Qiagen: Mississauga, Ontario, Canada)

PCR

Expand high fidelity PLUS PCR system, dNTPack Cat. No. 04 743 733 001 (Roche Diagnostics: Mannheim, Germany)

Custom primers (Invitrogen: Burlington, Ontario, Canada)

Protein quantification and preparation

BCA protein assay kit, Cat. No. 71285-3 (Novagen: Gibbstown, New Jersey, U.S.A.)

Sequencing

ABI prism BigDye terminator cycle sequencing ready reaction kit v.3.1 (Applied Biosystems, Streetsville, Ontario)

BigDye[®] 3.1 terminator cycle sequencing kit (Applied Biosystems, Foster City, California, U.S.A.)

Western blotting

iBlot[®] transfer stack (nitrocellulose) Cat. No. IB3010-01 (Invitrogen: Burlington, Ontario, Canada)

Novex[®] ECL chemiluminescent substrate reagent kit, Cat. No. WP20005 (Invitrogen: Burlington, Ontario, Canada)

3.1.4 Specialized equipment

DNA extraction

BioRobot EZ1 (Qiagen: Mississauga, Ontario, Canada)

Nucleic acid quantification

NanoDrop 1000 spectrophotometer (Thermo Scientific: Wilmington, Delaware, U.S.A.)

Agilent BioAnalyzer 2100, Part No. G2938C (Agilent: Santa Clara, California, U.S.A.)

PCR

Thermocycler (MJ Research: Waltham, Massachusetts, U.S.A.)

Western blotting

iBlot[®] gel transfer device, Cat. No. IB1001 (Invitrogen: Burlington, Ontario, Canada)

Molecular imager ChemiDoc XRS+ system, Cat. No. 170-8252 (Bio-Rad: Mississauga, Ontario, Canada)

ChemiDoc universal hood, Cat. No. 170-8126 (Bio-Rad: Mississauga, Ontario, Canada)

Sequencing

ABI prism 3100 genetic analyzer (Hitachi, Japan)

3.1.5 Software programs

GraphPad Prism 4.0 (GraphPad Software: San Diego, California, U.S.A.)

MEGA version 4.0: Molecular Evolutionary Genetic Analysis (Biodesign Institute: Tempe, Arizona, U.S.A.)

Hypermut (<http://www.hiv.lanl.gov/content/sequence/HYPERMUT/hypermut.html>)

Recombination Identification Program 3.0 (RIP 3.0)

(<http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html>)

Sequencher v.4.2 (Gene Codes Corporation: Ann Arbor, Michigan, U.S.A.)

3.2 Study population

3.2.1 Pumwani CSW cohort

The CSW cohort was first established in the Pumwani slum of Nairobi, Kenya in 1985 (155). The CSW cohort is comprised of women who are actively involved in sex work (54, 55). The sex worker cohort enrollees complete an administered, self-reported survey, which describes their sexual practices and other epidemiological factors each year. All individuals in this study were part of a well described cohort of commercial sex workers from the ML clinic (54). All experiments were conducted with each participant's understanding and consent. The work conducted from experiments in this thesis are from subjects who are HIV⁺, ARV naïve at the time samples of blood were collected and acquired HIV-1 infection through heterosexual contact. Ethical approval was obtained from both the University of Manitoba and University of Nairobi ethics review committees to conduct these studies.

3.2.2 Subject selection for APOBEC3G degradation

HIV-1 proviral DNA was isolated from a total of six subjects, three of which who demonstrated to have high levels of APOBEC-mediated HIV-1 proviral hypermutation in the vpu/env region of proviral HIV-1 genome (MLs 1592, 1970, and 1975) and three subjects of who did not show as high levels of APOBEC-mediated HIV-1 proviral

hypermutation (ML 1418, 1481, and 1868) in the *vpu/env* gene. Vif from these subjects were chosen to test the degradation of APOBEC3G in an *in vitro* system.

3.2.3 Subject selection for sequencing: examination of HIV-1 proviral hypermutation at the *pol* gene

Based on random sample availability, a total number of five hundred fifteen whole blood and PBMC samples were collected (2007-2009) from patients enrolled in the ML clinic. All five hundred fifteen HIV-1 positive, antiretroviral naïve women were believed to be infected with HIV-1 by heterosexual contact at the time blood samples were collected. The following experiments were conducted using DNA obtained from members of this cohort with informed consent.

3.3 Protocols

3.3.1. Extraction of total genomic DNA

DNA was extracted from women in the ML cohort by two methods: either from PBMCs using a Qiagen DNA extraction kit or DNA was extracted directly from the whole blood using a Qiagen EZ1[®] Kit and BioRobot EZ1. To determine if DNA was suitable for conducting PCR experiments, the quality of DNA was assessed by measuring the concentration of DNA (50ng/ul) using the Nanodrop1000 spectrophotometer.

3.3.2 Amplification of HIV-1 provirus

3.3.2.1 Full length *pol* amplification

Pol was amplified by nested PCR reaction which generated a 3148 bp amplicon in

the first nested PCR reaction and the second nested PCR reaction generated a 1942 base pair (bp) amplicon. Two different and specific primer groups were used in the first nested PCR reactions to determine which primer group yielded favourable results by visualizing the PCR products from the two different primer groups on a 1% agarose gel. The first primer group of the first nested PCR reaction were: Gagseqf4b and Polseqr2 with the alternate primer group of the first nested reaction were: Polseqf0 and MM4bb/Polseqr2.5. The first primer group was chosen to subsequently amplify the pol region of HIV-1 genome as the second primer group did not work as well as the first primer group. The first primer group was chosen over the second primer group as the second primer group did not show up well on the 1% agarose gel suggesting the second primer group did not amplify the DNA well. The specific primers used in the second nested PCR reactions are: Polseqf1 and NPol4481b/Polseqr3. Table 1 lists the following primers used in amplifying pol along with their sequence orientation, location, melting temperature (T_m), and annealing temperature (AT).

Generation of a 3148 bp amplicon in the first nested PCR reaction utilized the following conditions (50 µl final volume/reaction): 10 µl 5x buffer, 1 µl deoxynucleotide triphosphates (dNTPs), 1 µl of each forward and reverse primer, 0.5 µl enzyme, 33.5 µl water (H₂O) and 3.0 µl DNA sample. To generate a 1942 bp amplicon in the second nested PCR reaction consisted of the following 50 µl reaction: 10 µl 5x buffer, 1 µl dNTPs, 1 µl of each forward and reverse primer, 0.5 µl enzyme, 36.5 µl H₂O, and 2 µl DNA from the first run nested PCR reaction. PCR thermocycling conditions for each of the first and second nested PCR reactions are described in Table 2. The amplicons were examined by visualizing the second round PCR products after gel electrophoresis on a

Table 1: Pol primers used for nested-PCR, their gene locations, melting and annealing temperatures

Primer Name	Primer Sequence	Orientation	Gene Location^a	Tm (°C)	AT (°C)
Gagseqf4b ^b	5'TTAGATACAGGAG CAGATGATACAG3'	Forward	2322-2346	58.7	52.0
Polseqr2 ^c	5'ATCCTACCTTGT TATGTCCTG3'	Reverse	5470-5450	56.1	52.0
Polseqf0 ^c	5'GYAYAAAYAATG AGACACCAG3'	Forward	2950-2970	52.2	48.0
MM4bb/ Polseqr2.5 ^d	5'TGGATGTGTA CTGAACTTA3'	Reverse	5213-5192	54.5	48.0
Polseqf1 ^c	5'TATCAGTACAATG TGCTTCCAC3'	Forward	2979-3000	56.3	53.0
NPol4481b/ Polseqr3 ^d	5'CTGCTGTCCCTGT AATAAACCCG3'	Reverse	4921-4899	62.0	53.0

^aGene location is given in HXB2 numbering.

^bPreviously designed laboratory primers.

^cPrimers designed by Dr. Allison M. Land.

^dPrimers designed by Dr. Melanie Murray.

Table 2: Thermocycler conditions for nested-PCR amplification

Step	Temperature (°C)	Time (sec)	Cycle
Thermocycling Run No. 1			
1. Initialization	94	120	1
2. Denaturation	94	15	1
3. Annealing	52	30	1
4. Elongation	72	120	1
5. Repeat steps 2 to 4 for 30 cycles			
6. Final Elongation	72	420	1
7. Final Hold	4	Indefinite	--
Thermocycling Run No. 2			
1. Initialization	94	120	1
2. Denaturation	94	15	1
3. Annealing	53	30	1
4. Elongation	72	90	1
5. Repeat steps 2 to 4 for 25 cycles			
6. Final Elongation	72	300	1
7. Final Hold	4	Indefinite	--

transilluminator, second round PCR products were purified and sequenced. The detailed steps will be discussed in sections below.

3.3.2.2 Partial vpu/env region amplification

Patient samples observed to have significant hypermutation in the pol region of HIV-1 were selected for vpu/env amplification to determine if there is proviral hypermutation in both the pol and vpu/env genome of HIV-1. Nested PCR reactions amplified proviral vpu and the first 349 nucleotides of env. The first nested PCR reactions generated a 2850 bp amplicon. The primers, Polseqf3.6 and Envseqr4.5, were used for the first nested PCR reactions. Envpcrf and Envseqr6 were used in the second reaction and generated a 646 bp amplicon. Table 3 lists the following primers used in amplifying vpu/env along with their sequence orientation, location, T_m, and AT. The exact thermocycling conditions were followed as the full length pol amplification and this can be found in section 3.3.2.1 of this thesis

3.3.3 Visualizing PCR products by gel electrophoresis

The second round nested PCR products were checked to determine if the reactions were successful. A total of 5 µl DNA from the second round nested PCR reaction and 0.8 µl 6x loading dye were mixed together and loaded onto a 1% agarose DNA gel containing EtBr (concentration 10 mg/mL) submerged in TBE buffer. The gel was resolved 1 hour at 90 volts (V) and DNA fragments were visualized under ultraviolet light on a transilluminator and photographed to check for the correct sized PCR products after amplification.

Table 3: Vpu/env primers used for nested-PCR, their gene locations, melting and annealing temperatures

Primer Name	Primer Sequence	Orientation	Gene Location^a	Tm (°C)	AT (°C)
Polseqf3.6 ^c	5'AGTTATCCCAGC AGAAACAGGAC3'	Forward	4490-4512	60.2	52.0
Envseqr4.5 ^c	5'TGTTATTTCTAGA TCCCCTCCTG3'	Reverse	7340-7318	58.4	52.0
Envpcrf ^c	5'GGCTTAGGCATCTCC TATGGCAGGAAGAAG3'	Forward	5954-5983	67.5	52.0
Envseqr6 ^b	5'CGAGTGGGGTTAA CTTACACATG3'	Reverse	6600-6577	60.3	52.0

^aGene location is given in HXB2 numbering.

^bPreviously designed laboratory primers.

^cPrimers designed by Dr. Allison M. Land.

3.3.4 PCR product purification

The second round PCR reactions that had a single amplicon of the correct size were either purified kindly by the National Microbiology Laboratory (NML) DNACore Facility using a robot or using a Montage PCR centrifugal filter device (Millipore: Billerica, Massachusetts, Cat No. P36461) containing two filtrate collection vials per PCR reaction as described by the manufacturer. A modification was made to the user guide as a smaller volume of PCR products was used for experiments. The sample reservoir was adjusted to a final volume of 400 μ l which contained 355 μ l TE buffer and 45 μ l PCR DNA per reservoir (thirteen products were purified by this method).

3.3.5 DNA sequencing

After DNA purification, sequencing was set up as essentially described in instructions for the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit. Each reaction consisted of: 1.75 μ l sequencing buffer, 0.5 μ l of Big Dye[®] v3.1, 1.5 μ l sequencing primer (concentration 10 ng/ μ l), 3.25 μ l H₂O, and 3.0 μ l of purified template PCR product. The sequencing reactions were resolved on an ABI Prism 3100 Genetic Analyser and the resulting trace files were assembled into overlapping, ds contigs, using Sequencher v.4.0.5 for the analyzes of two hundred forty-one sequences.

3.3.5.1 Full length pol sequencing

Primers for sequencing 2 kb of the HIV-1 *pol* gene required 3 sequencing reactions (two primers in both directions/reaction) to cover the entire length of *pol*. The

first sequencing reaction consisted of the forward and reverse primers, Polseqf1.5 and MM3bb/Polseqr4.5, respectively. The second sequencing reaction consisted of the forward and reverse primers, Polseqf2.5 and POL4R, respectively. And the third sequencing reaction consisted of the forward and reverse primers, Polseqf3.6 and Seq3630b/Polseqr5.5, respectively. The specific primers designed across the length of pol are outlined in Table 4 along with their sequence orientation, location, T_m, and AT.

3.3.5.2 Partial HIV-1 proviral genome sequencing

The partial proviral pol segment was sequenced using MM3ab/Polseqf3.5 and POL4R. The vpu/env segment was sequenced using the nested PCR primers, Envpcrf and Envseqr6. Please refer to Table 4 for their orientation, location, T_m, and AT.

3.3.6 HIV-1 sequences and phylogenetic analysis

HIV-1 sequences from each subject (full-length or partial) were aligned against the current reference sequences (2007) from the Los Alamos HIV (www.hiv.lanl.gov) sequence database using ClustalW. MEGA 4.0, a phylogenetic analysis program, was employed for neighbour-joining phylogenetic tree analysis (nucleotide distance calculated by Kimura's two parameter method) to determine what subtype each ML patient's HIV-1 sequence belonged to. The strength of the branch nodes was tested with a bootstrap analysis of 500 replicates. Sequences were further analysed online using the Recombination Identification Program from the Los Alamos Sequence database to determine if sequences contained any recombinants.

Table 4: Sequencing primers for pol (full-length and partial) and their gene locations, melting and annealing temperatures

Primer Name	Primer Sequence	Orientation	Gene Location^a	T_m (°C)	AT (°C)
Polseqf1.5	5'GATGAYTTRTATG TAGGATCTG3'	Forward	3102-3123	52.6	53.0
MM3bb/ Polseqr4.5	5'CCTTTGTGTGCTG GTACCCATG3'	Reverse	4171-4150	61.9	56.0
Polseqf2.5	5'GAYRGA CTAYTGG CAGGCTAC3'	Forward	3755-3775	58.0	53.0
POL4R	5'GCTGTCCCTGTAA TAAACCCG3'	Reverse	4919-4899	60.0	55.0
Polseqf3.6	5'AGTTATCCCAGCA GAAACAGGAC3'	Forward	4490-4512	60.2	56.0
Seq3630b/ Polseqr5.5	5'ATTGTTTTACATC ATTAGTGTG3'	Reverse	3651-3630	50.8	55.0

^aGene location is given in HXB2 numbering.

^bPreviously designed laboratory primers.

^cPrimers designed by Dr. Allison M. Land.

^dPrimers designed by Dr. Melanie Murray.

3.3.7 Detection of hypermutation at HIV-1 proviral genome

HIV-1 sequences from each subject (full-length or partial) were analyzed in a program, Hypermut, to detect APOBEC-induced hypermutation from the Los Alamos sequence database. Each sequence was compared on a subtype by subtype basis to a clade specific consensus sequence.

3.3.8 Statistics used for data analysis

For this thesis, two statistical tests were employed: the Mann-Whitney test and Chi-Square test which were both determined using GraphPad Prism 4.0 software. We considered a p-value of less than 0.05 significant.

3.3.9 Cell culture

The human cell line, 293T cells, derived from human embryonic kidney (HEK) cells were cultured in DMEM supplemented with 10% FCS, 1x penicillin/streptomycin and grown at 37°C, 5% CO₂ and 100% relative humidity (REVCO ULTIMA: RC03000D-7-ABB).

3.3.10 Plasmids

The plasmid constructions, SCMV-IN and pCMV-HA (with APOBEC3G insert), were kindly donated by Dr. Xiaojian Yao (Figure 10 and Figure 11). Appropriate vif sequences were subcloned into a multiple cloning site (MCS) of the SCMV-IN expression plasmid by Dr. Allison Land and the sequence of each patient were confirmed to be the

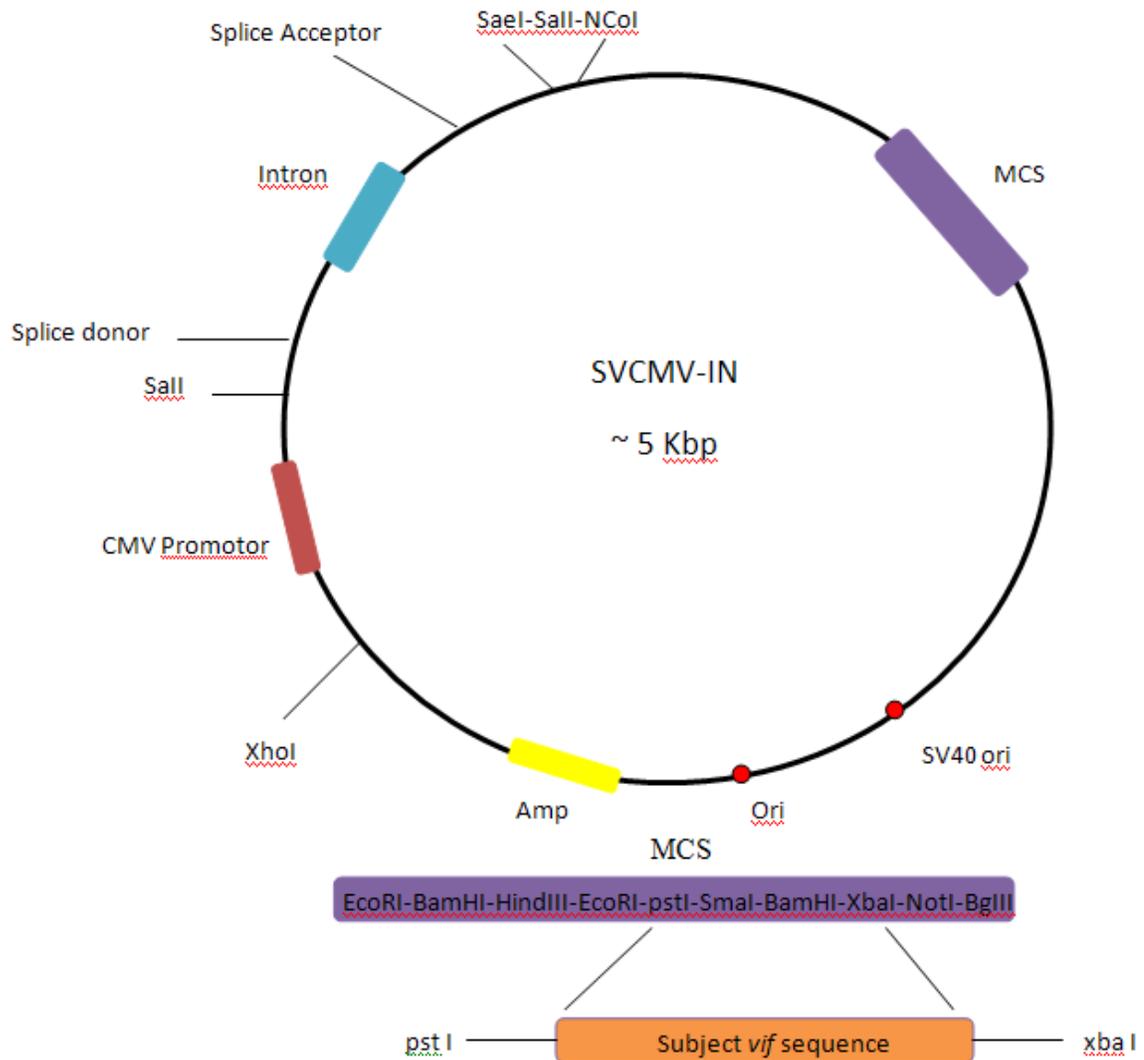


Figure 10: Schematic of SVCMV-IN expression vector encoding several markers. Important markers are: human cytomegalovirus immediate early promoter/enhancer (CMV promoter), origin of replication (ori), ampicillin resistance (amp) marker for selection in *E. coli*, and multiple cloning site (MCS). Subject vifs were cloned into pst I and xba I MCS. Verification of insert can be determined using xba I and pst I. Vectors with *vif* inserts were kindly obtained from Dr. Allison Land.

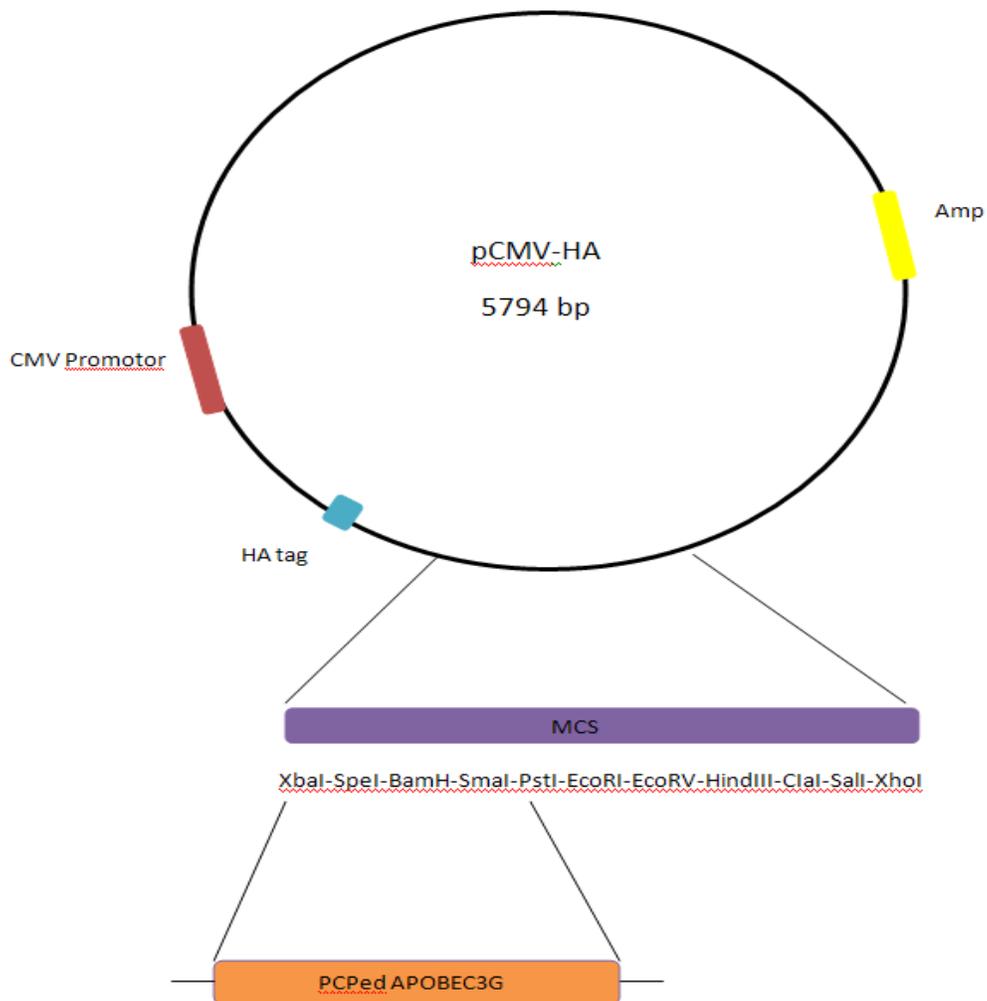


Figure 11: Schematic of pCMV-HA mammalian expression vectors encoding several markers. Important markers are: human cytomegalovirus immediate early promoter/enhancer (CMV promoter), haemagglutinin (HA) epitope tag, multiple cloning site (MCS), and ampicillin resistance (amp) marker for selection in *E. coli*. Verification of insert can be determined using xba I and pst I. Plasmid product no. PT3283-5 obtained from Clontech Laboratories, Inc.

correct vif insert by Dr. Allison Land and expression plasmids were stored in glycerol stocks at -80°C.

3.3.11 Plasmid DNA amplification, purification and quantification

Two different kits were tested to determine which method would yield the highest plasmid DNA. When the first kit was tested, QIAprep miniprep kit (250), according to manufacturer's instructions this resulted in low quantities of plasmid DNA; despite the fact that all plasmids used were known to be high-copy number plasmids. Several modifications were made to the original QIAprep miniprep kit protocol to optimize plasmid DNA yield. Protocol 1 followed the manufacturer's protocol where it yielded the lowest concentration of plasmid as determined by utilizing a spectrophotometer instrument (NanoDrop™ 1000: Thermo Scientific). Slight modifications were made (called protocol 2) by doubling the quantity of bacterial cells used to isolate and purify plasmid DNA and this slightly increased the yield. An alternate kit was examined, the QIAprep maxiprep kit (protocol 3) from Qiagen and compared to protocol 1 and protocol 2, there was an observable increase in plasmid DNA. However, protocol 2 gave the best results since it was more consistent and used a smaller volume of reagents than protocol 3, it was further optimized (called protocol 4) by increasing the incubation times when buffers were added. Protocol 4 was the method of choice for all future plasmid DNA purification and quantification. The differences in protocols can be observed in Figure 12.

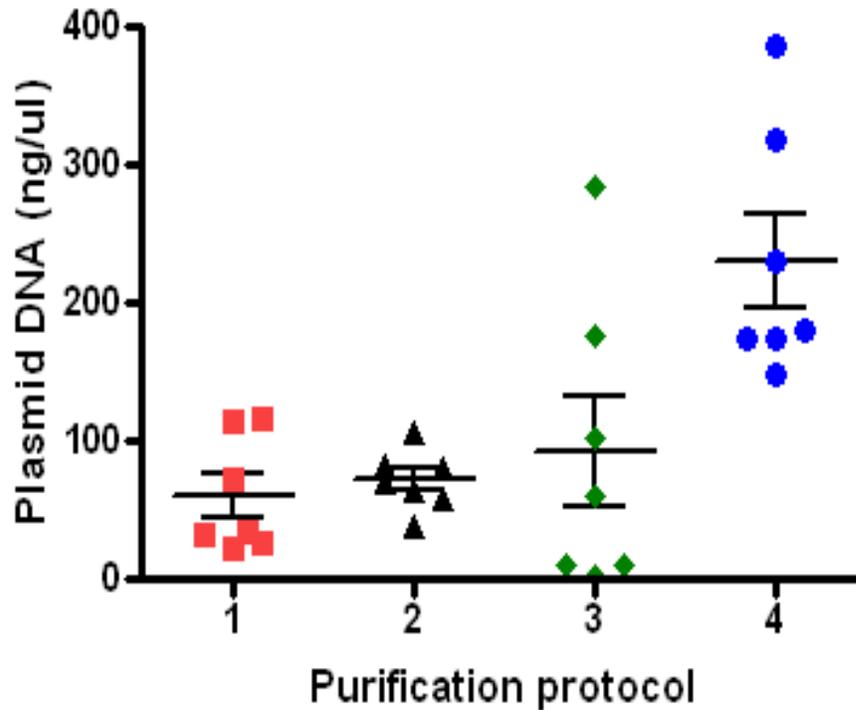


Figure 12: Plasmid DNA yields from four different protocols. Several protocol variations were attempted to maximize plasmid DNA yields from *E. coli* strain DH5-*a* cells. Protocol 1, the original method followed from QIAprep miniprep kit by Qiagen. Protocol 2 was similar to protocol 1 by increasing the number of bacterial cells. Protocol 3 used a different kit from Qiagen called the QIAprep maxiprep kit and was not favorable due to the large volume of materials used to carry out the experiment. Protocol 4 included slight modifications to protocol 2 by further increasing bacterial cell concentrations and altering incubation times of buffers

3.3.12 Restriction enzyme digests of vectors

The purified plasmids were examined to determine if the successful fragments were inserted by performing a restriction digest using the restriction enzymes pst I and xba I. Each reaction contained 12 μ l H₂O, 10 μ l 10x React[®] 2 buffer, 0.5 μ l pst I, 0.5 μ l xba I, and 5 μ l purified plasmid. Each reaction mix was vortexed, quick-spinned and incubated at 37°C for 1.5 hours. After incubation, the digestion products were resolved on a 1% agarose gel containing EtBr to check for correct sized inserts. The gel ran for 22 minutes at 126 V and the gel was placed on a transilluminator where it was photographed, and the correct sized inserts were confirmed.

3.3.13 Transfections of 293 T cells

HEK 293 T cells were plated into 6 well plates (concentration of 0.5×10^6 cells/well) 24 hours prior to transfection. On the day of experiment, media was gently removed from all wells and 1 mL of fresh warm media was added 30 minutes prior to transfection. Sixty to eighty percent confluent cells were transfected with 2 μ g of DNA of the SVCMV-IN expression plasmid containing the appropriate vif inserts and 1 μ g DNA of pCMV-HA expression plasmid containing the APOBEC3G insert (ratio 2:1) using the Mirus TransIt-LT1 reagent. Media was carefully removed prior to addition of transfection/DNA mixture. Two microlitres of transfection reagent per μ g DNA was added to 250 μ l DMEM (Sigma), gently mixed, and incubated at room temperature for 30 minutes. The transfection mixture was added in a drop-wise manner to the cells followed by 30 minute incubation at 37°C, 5% CO₂. Cells were covered with 2 mL media and

incubated for 48 hours before harvesting protein.

3.3.14 Preparation of cell lysates

Transfected cells were collected on ice by detaching the cells from the well by washing with the media (2 ml/well) present in each well. Cells were collected into sterile 1.5 mL eppendorf tubes. Media and cells were centrifuged at 4000 rpm for 5 minutes at 4°C and supernatants were discarded. Before discarding transfection plates, a final 0.5 mL cold PBS wash was conducted by quickly rinsing off any cells remaining in the well and then transferred to the appropriate corresponding eppendorf tubes. Tubes were centrifuged again at 4000 rpm for 5 minutes at 4°C and supernatants were removed. A quick wash was performed by gently resuspending pellets into 1 mL of fresh PBS and then centrifuged at 4000 rpm for 5 minutes. PBS was carefully removed from pellet using a p200 pipette. Then 100 µl RIPA lysis buffer was added to each pellet and pipetted vigorously to lyse cells. Cell lysates were incubated for 30 minutes on a rotator at 4°C and then centrifuged at 10000 rpm for 15 minutes at 4°C. Supernatants were transferred to clean eppendorf tubes and stored at -25°C.

Frozen PBMCs were removed from liquid nitrogen and vials were suspended in a 37°C water bath until cells were slightly thawed. One mL of RPMI supplemented with 10% FCS and 1x penicillin/streptomycin was then added to each vial to prevent cells from undergoing shock before being resuspended into 8 mL RPMI supplemented media. Cells were centrifuged for seven minutes at 1400 rpm and supernatants were discarded. Resuspended PBMCs were counted with a haemocytometer and then cells were lysed by the addition of 100 µl RIPA lysis buffer for every 1×10^6 cells and soluble supernatants

were collected as above.

3.3.15 Quantification of protein concentration in cell lysates

Protein concentration was determined using the BCA protein assay kit which contained two reagents, BCA solution and 4% cupric sulfate, and BSA as a standard at a concentration of 2 mg/mL. This kit was used to determine protein concentrations present in cell lysates before Western blotting. According to the manufacturer's instructions, BSA was prepared in the diluents used to lyse cells (RIPA lysis buffer) over the range of 0-1000 µg/mL in sterile eppendorf tubes. To be certain that protein concentration would be within the detection limit of the assay, each protein lysate was diluted 1:10 in RIPA lysis buffer (in duplicate). The working reagent was prepared by combining 200 µl of BCA solution with 4 µl of 4% cupric sulfate per sample.

Twenty-five µl of diluted samples and standard were transferred to an ELISA plate. Two hundred microlitres of working reagent was added to each sample, shaken briefly on a rocker (Bellco: Rocker Platform 7740-10010) and incubated at 37°C for 30 minutes. After incubation, samples and standards were read at 562 nm on a microplate reader (Molecular Devices 02500: SPECTRA max PLUS 384). Samples were done in duplicate to ensure the accuracy of the assay. Standard curves were generated based on their absorbance readings and a linear graph was made while the protein concentration of each sample was determined by an equation generated by the standard curves using a microplate data acquisition and analysis software package (SoftMax Pro: Molecular Devices).

3.3.16 Western blots

Twelve µg of each protein sample were loaded onto a 4-12 % NuPAGE[®] SDS-PAGE gel and electrophoresed at 200 V for approximately 40 minutes using a gel electrophoresis unit (XCell4 SureLock[™] Midi-Cell Runner ©: Invitrogen). The gel was rinsed with ddH₂O, and proteins were transferred onto a nitrocellulose membrane using the Invitrogen iBlot[®] gel transfer device according to the manufacturer's instructions. The membrane was blocked using 5% milk blocking solution for 1 hour at room temperature. Ten mL of a 1:1000 dilution of primary antibodies specific for HIV-1 Vif and APOBEC3G were applied to each blot, incubated overnight at 4°C. Membranes were washed with 10 mL 1x TBST for five minutes and washing was repeated three more times. Membranes were then incubated with 10 mL of a 1:5000 dilution of a goat anti-rabbit secondary antibody conjugated with horse-radish peroxidase for one hour and then washed with 10 mL 1x TBST for five minutes (repeated four times in total). Membranes were incubated with Novex[®] ECL chemiluminescent substrate reagent kit as per manufacturer's instruction and ECL bands were determined using the Quantity One[®] software package (Bio-Rad) for imaging and analyzing 1-D electrophoresis gels.

4.0 Results

4.1 HIV-1 proviral *pol* gene sequence reflects APOBEC-mediated hypermutation

4.1.1 Defining hypermutated sequences and HIV-1 subtype assignment

To test the hypothesis that APOBEC-mediated hypermutation associates with reduced HIV-1 disease progression, we first examined the region of HIV-1 proviral *pol* that theoretically should be most affected by APOBEC3G/F activity (189). Two-hundred forty one proviral HIV-1 *pol* sequences (sequenced in both directions and assembled in the program Sequencer) were examined from HAART naïve, HIV⁺ women in the CSW cohort. Before sequences were analyzed for hypermutation, they were imported into a Recombination Identification Program 3.0 (RIP 3.0) to test for recombination. RIP 3.0 is a program used for HIV-1 subtyping and consists of consensus sequences for near full-length subtypes A1, B, C, D, F1, F2, G, H, and CRF01 HIV-1 genome. We confirmed thirteen recombinant sequences and further we were able to determine the subtype of our *pol* sequences by analyzing the graphs generated. The majority of sequences from this study were from subtype A1 (n=161) while subtype D (n=46) were the next highest number of sequences. The orders of the other subtypes were as follows: C (n=18), recombinants (n=13), and G (n=3). Within the recombinants, the sequences consisted of the following mixtures of subtypes: A1_C (n=1), C_D (n=2), A1_D (n=3), A1_B_D (n=1), A_B_C_D (n=1), and A_B_C_F (n=1). The last four sequences were either CRF's and/or CRF's mixed with subtypes which consisted of the following: A2.CY.94CY017-41 (n=2), A2.CY.94CY017-41_G (n=1), and A2.CY.94CY017-41_D (n=1). These results give us an understanding as to what subtype of the virus is infecting each subject in this

study. The percentages of subtype and recombination sequences are described in Figure 13.

The HIV-1 pol sequences were also analyzed by phylogenetic analysis using Molecular Evolutionary Genetic Analysis, version 4.0 (MEGA 4.0). Sequences were subjected to neighbour-joining phylogenetic tree analysis (nucleotide distance calculated by Kimura's two parameter method) of the full-length pol sequences (n = 28) as a way to visually observe which subtype each sequence belonged (Figure 14). The strength of the branch nodes was tested with a bootstrap analysis of 500 replicates. A list of subjects and their subtype used for full-length pol sequence analysis (n = 28) can be found in Table 5. A list of subjects used for partial pol sequencing (n = 211) will be shown in a later section of this thesis.

To determine the extent of HIV-1 hypermutation among subjects, we examined a total of 241 pol sequences encompassing the last 705 bases of pol. Some sequences were obtained from two different time points from a single subject to assess the potential stability of APOBEC-mediated hypermutation. Sequences from each subject were grouped and aligned together by subtype into MEGA 4.0 by ClustalW analysis. Using the web tool Hypermut 2.0 from the Los Alamos database that detects the physical location of each mutation in a subject, we generated a mutation map (Figure 15). This allows the visualization of the total number of substitutions occurring in a subject for a particular subtype when compared to the consensus sequence. The consensus reference sequences generated from each subtype was obtained from the Los Alamos HIV sequence database. Based on the graphs generated, each subject had varying levels of hypermutations at different points along the proviral HIV-1 pol genome.

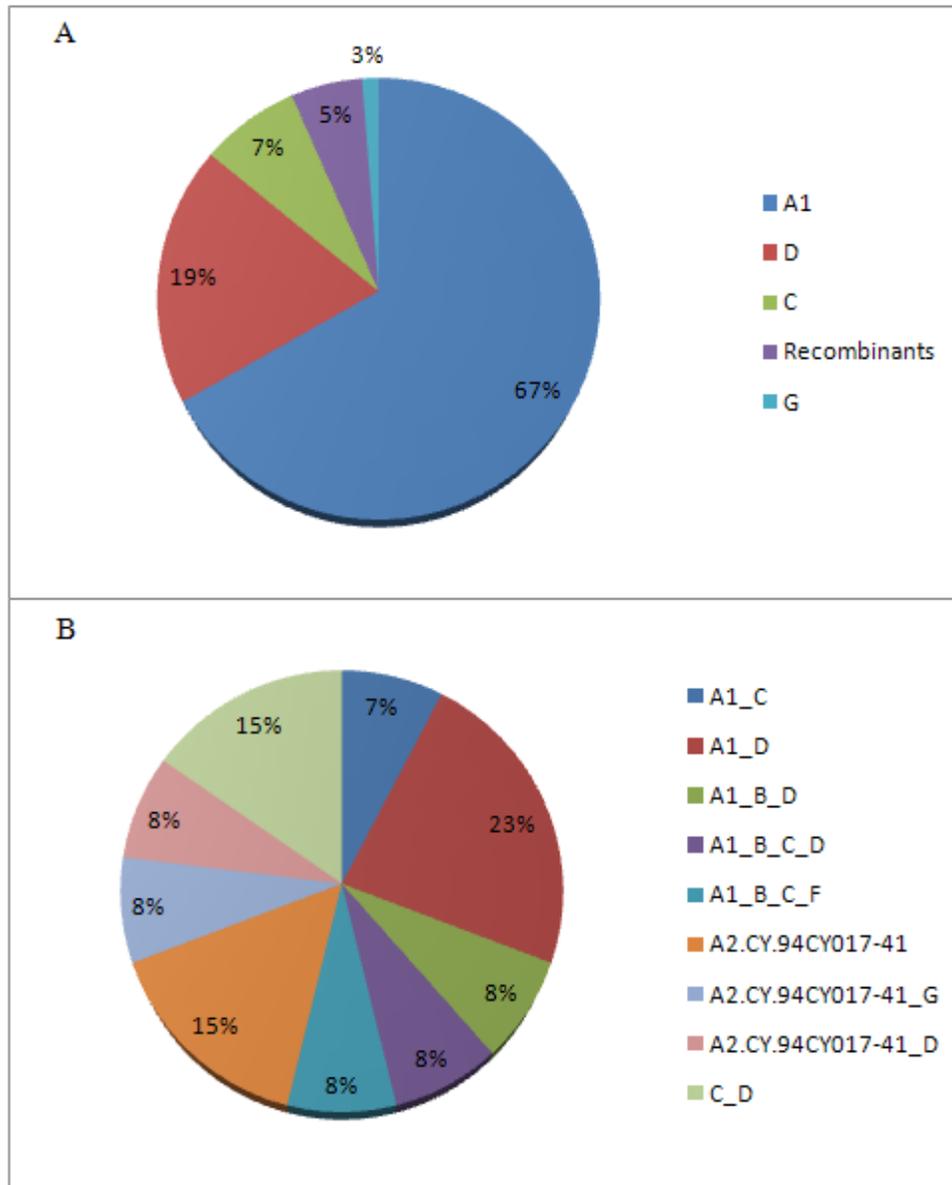


Figure 13: Relative proportions of HIV-1 subtypes and recombinants for 241 HIV-1 proviral pol sequences. (A) The percentage of the 241 pol sequences organized into a pie chart with the appropriate subtypes and recombinants analyzed in this thesis. (B) The percentage of 13 pol recombinant sequences displayed in a pie chart. Colored legend indicates what percentage of subtypes and recombinants are in pol sequence pool.

Table 5: HIV-1 subtype classifications from twenty-eight full length pol sequences generated in this thesis

<u>Subject</u>	<u>Subtype</u>	<u>Subject</u>	<u>Subtype</u>
ML2503	A1	ML2569	D
ML2522	D	ML2624	C
ML2525	D	ML2651	C
ML2560	A1	ML2713	C
ML2584	A-B-C-D	ML2719	A1-D
ML2587	A1	ML2460	A1
ML2594	A1	ML2468	A1
ML2614	A-B-D	ML2604	C
ML2645	A1	ML2630	A1-C
ML2700	D	ML2655	A1
ML2704	A1	ML2667	A1
ML2715	A1	ML2683	A1
ML2750	A-D	ML2709	C
ML2469	A1	ML2751	D

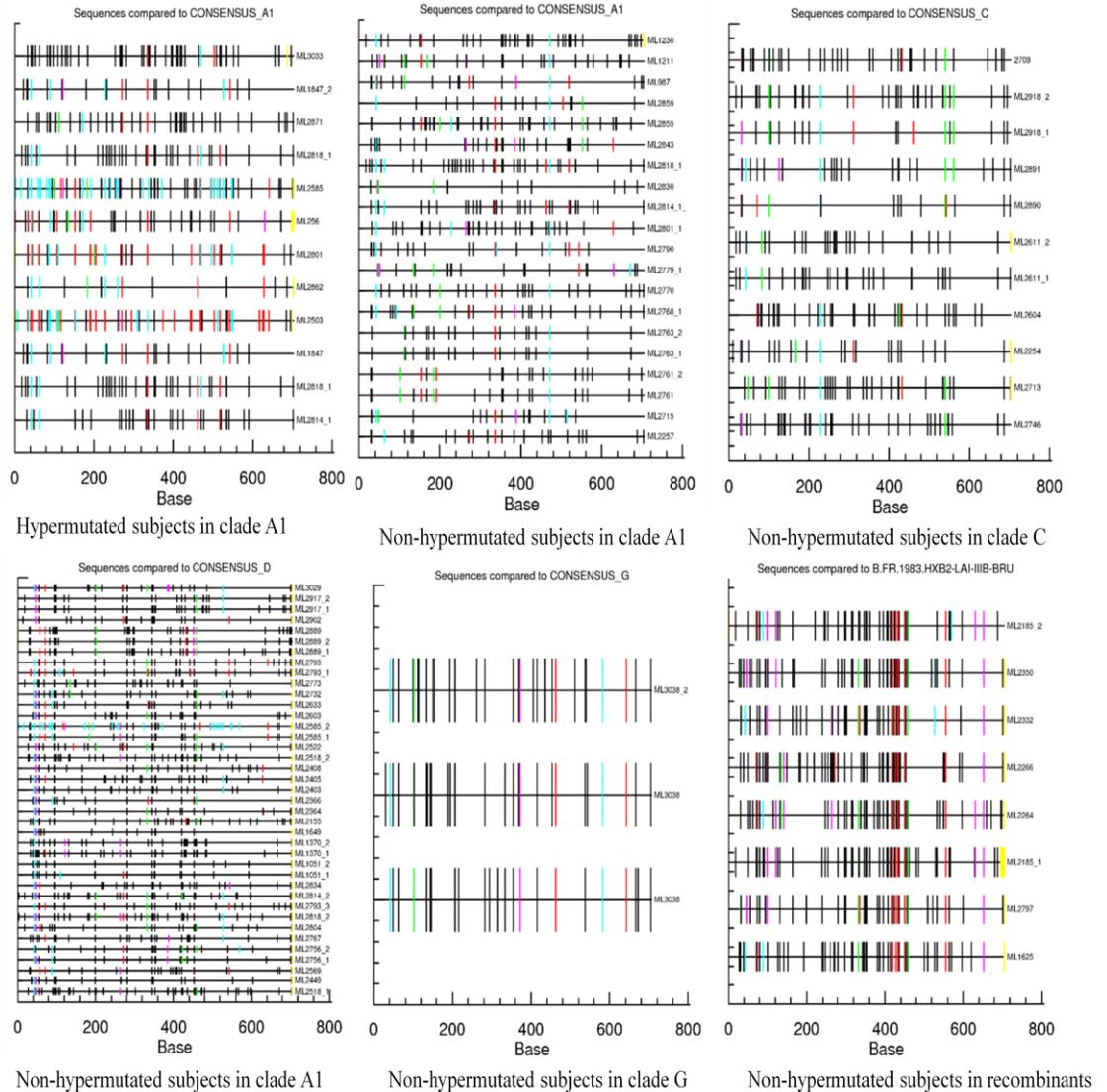


Figure 15: Graphical representation of mutations detected by using the tool hypermut 2.0. Each sequence from a patient must be grouped by its subtype type and compared to a consensus sequence to determine mutations in a single subject. Each colored mark along a single line represents a mutation in one subject. All colored marks represent G to A changes (red = GG to AG, cyan = GA to AA, green = GC to AC, and magenta = GT to AT). Black represents no G to A changes and yellow indicates nucleotide gaps in the sequences.

4.1.2 Sequencing data reveal APOBEC3G/F-mediated hypermutation at HIV-1 proviral *pol* gene

Using Hypermut 2.0, a total of sixteen out of two hundred forty-one *pol* sequences, or 7% of examined sequences, were hypermutated (Table 6). Hypermutation by APOBEC3G/F is present when specific AA and AG nucleotide changes occur more frequently than random mutations would occur, this is defined as hypermutation ($p < 0.05$). Two of the sixteen hypermutated *pol* sequences were identified by full length *pol* sequencing. As shown in Table 6, the sixteen proviral hypermutated *pol* sequences showed significant hypermutation activity by APOBEC3G, APOBEC3F, or a combination of both. The majority of subjects with detectable hypermutation (except for the two ML1847 time points, ML2585, and ML2862) demonstrated APOBEC3G-mediated hypermutation (nucleotide changes from GG to GA in the plus strand of DNA) in the *pol* region of proviral HIV-1. The four samples (actually, only 3 subjects) (both ML1847 time points, ML2585, and ML2862) that did not have detectable APOBEC3G activity in the *pol* region demonstrated APOBEC3F activity (nucleotide changes from AG to AA in the plus strand of DNA). Eight of the sixteen samples with hypermutated proviruses had evidence of both APOBEC3G and APOBEC3F enzymatic activity (the two ML2818 time points, ML2503, ML2801, ML2766, ML2871, ML2503, and ML2645). Tables 7, 8, 9, 10, and 11 show all non-significant hypermutated sequences from *pol* sequencing grouped appropriately into subtypes: A1, C, D, G, and recombinants respectively.

Table 6: Sixteen samples in the pol region (nucleotide positions 4150-4919) demonstrated significant hypermutated HIV-1 provirus

Sample	Infection subtype	General hypermut significance ^b	Specific hypermut significance ^c	
			APOBEC3G	APOBEC3F
ML256	A1	0.0094	0.0014	<i>0.2419</i>
ML1847 ^a	A1	0.0307	<i>0.0916</i>	0.0485
ML1847 ^a	A1	0.0307	<i>0.0936</i>	0.0472
ML2503	A1	1.312 e ⁻⁰⁹	1.432e ⁻⁰⁹	1.068e ⁻⁰⁵
ML2503	A1	1.533e ⁻¹⁶	4.692e ⁻²⁰	1.331e ⁻⁰⁸
ML2585 ^a	D	8.224e ⁻⁰⁶	<i>0.4672</i>	1.211e ⁻⁰⁸
ML2645	A1	0.0013	2.295e ⁻⁰⁶	<i>0.2800</i>
ML2766	A1	2.972e ⁻⁰⁷	3.119e ⁻¹⁰	0.0069
ML2801 ^a	A1	1.327e ⁻⁰⁷	1.791e ⁻⁰⁹	0.0010
ML2814	A1	0.0322	0.0308	<i>0.1294</i>
ML2818	A1	0.0165	0.0299	0.0486
ML2818 ^a	A1	0.0190	0.0531	0.0510
ML2862	A1	0.0332	<i>0.0891</i>	0.0555
ML2871	A1	<i>0.0868</i>	0.0436	0.0407
ML2891	C1	<i>0.636</i>	0.0363	<i>0.5365</i>
ML3033 ^a	A1	<i>0.0632</i>	0.0317	<i>0.3543</i>

^aSuperscripts indicate sequences were obtained from two different time points during infection from a single subject to increase sample population of the study.

^bGeneral APOBEC-type hypermutation using Hypermut 2.0 analysis. Default settings of G to A substitutions with a downstream context of RD, where R is either A or G and D is A, T or G. Hypermutated sequences are defined as those with a p-value of <0.05.

^cHypermut 2.0 analysis for specific APOBEC3G (settings of G to A substitutions with a downstream context of GD) and/or APOBEC3F (settings of G to A substitutions with a downstream context of AD) hypermutation. Hypermutated sequences are defined as those with a p-value of <0.05.

Significant values are highlighted in yellow.

Non-significant values are italicized.

Subjects bolded are sequences generated from full-length genome of pol (nucleotide positions 3102-4919).

Within the significantly hypermutated sequences examined, we wanted to investigate the specific G to A changes that are caused by APOBEC3 proteins, APOBEC3G and APOBEC3F. In order to further characterize and investigate APOBEC3-associated HIV-1 hypermutations, we examined the dinucleotide sequence context for the sixteen subjects with proviral hypermutated sequences (Figure 16). Most of the G to A nucleotide sequence changes at pol were the result of the original nucleotide sequence being a GG or GA on the plus strand of DNA which suggest that there is APOBEC3G and APOBEC3F activity involved, as these nucleotide sequences change from GG to GA (demonstrates APOBEC3G activity) or AG to AA (demonstrates APOBEC3F activity). Based on the analyses of each subject, G to A nucleotide changes in the GG dinucleotide context (APOBEC3G activity) were more frequent than G to A changes in the GA context (APOBEC3F activity) within an individual, suggesting APOBEC3G is the predominant protein functioning in these particular subjects. In subject ML2891, however, the GA dinucleotide hypermutation is the only activity observed, suggesting that APOBEC3F is the most active editing enzyme in this subject. In all sixteen hypermutated subjects, very minimal or no changes occurs at the GT or GC dinucleotides meaning that these changes are not associated with APOBEC3G/F function. These findings suggest GG to GA (APOBEC3G activity) were the most frequently observed G to A changes followed by AG to AA (APOBEC3F activity). The low frequency observed in subjects illustrating GT to AT and GC to AC suggest they are due to non-APOBEC3G/F activity.

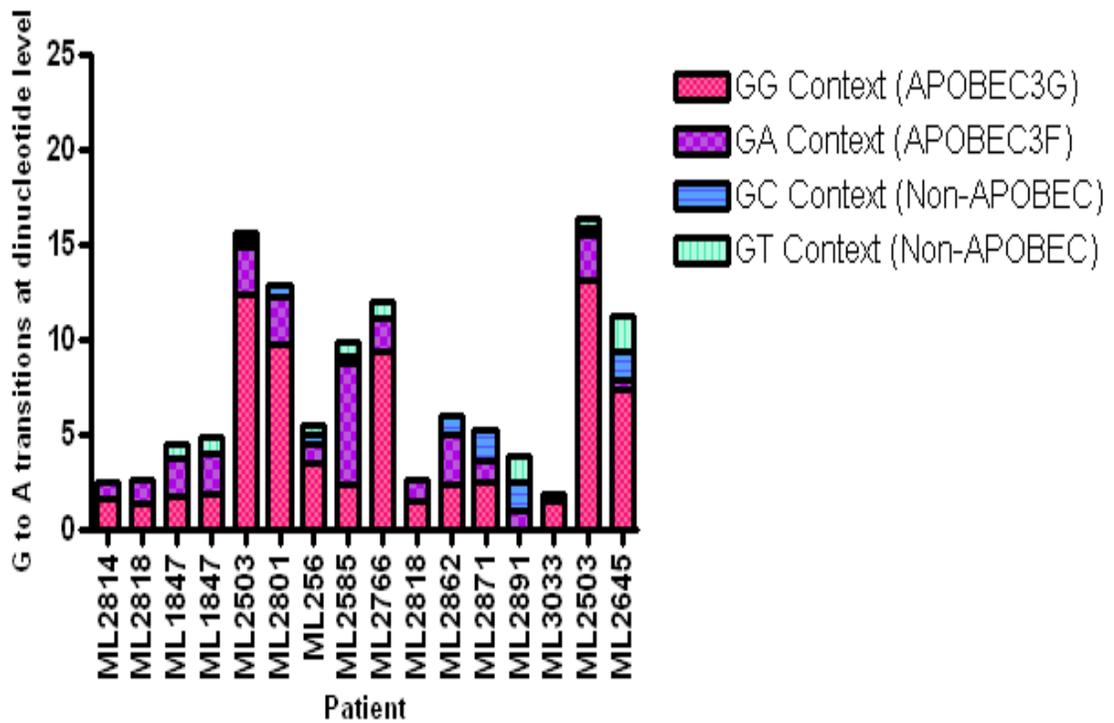


Figure 16: Histogram depicting dinucleotide context of G to A hypermutation from proviruses in sixteen subjects. The percentage of GG to AG, GA to AA, GC to AC, GT to AT occurring at the dinucleotide context is indicated. Hypermutation context for each patient was determined through comparison to a consensus sequence of its subtype, using Hypermut 2.0.

Table 7: Composition of one hundred thirty-four samples infected with HIV-1 subtype A1 with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)

<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>
ML0874	ML0987	ML1227	ML1211	ML1230	ML1230*
ML1266	ML1410	ML1430	ML1441	ML1529	ML1535
ML1575	ML1597*	ML1597*	ML1694	ML1725	ML1731
ML1740	ML1755	ML1778	ML1862	ML1875	ML1887*
ML1887*	ML2008	ML2014	ML2019	ML2039	ML2042
ML2042	ML2049	ML2052*	ML2052*	ML2102	ML2108
ML2126	ML2169	ML2181	ML2203	ML2218*	ML2218*
ML2219	ML2252	ML2257	ML2259	ML2297	ML2340*
ML2340*	ML2340*	ML2360	ML2359*	ML2359*	ML2391*
ML2391*	ML2393	ML2396*	ML2396*	ML2468	ML2469*
ML2469*	ML2487	ML2498	ML2542	ML2563	ML2566*
ML2566*	ML2587	ML2601*	ML2601*	ML2608*	ML2608*
ML2612	ML2614	ML2619	ML2622	ML2643	ML2645
ML2645	ML2682	ML2695	ML2715*	ML2715*	ML2719
ML2727*	ML2727*	ML2742	ML2761*	ML2761*	ML2763*
ML2763*	ML2766*	ML2766*	ML2768*	ML2768*	ML2770
ML2779*	ML2779*	ML2790	ML2793*	ML2801*	ML2830
ML2830*	ML2830*	ML2843	ML2845	ML2846	ML2846
ML2855	ML2859	ML2865	ML2871*	ML2871*	ML2877*
ML2877*	ML2895	ML2907*	ML2907*	ML2907	ML2933
ML3012*	ML3012	ML3012*	ML3013	ML3023*	ML3023
ML3023*	ML3026	ML3033*	ML3033	ML3034	ML3054
ML3054*	ML3054*				

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Table 8: Composition of twelve samples infected with HIV-1 subtype C with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)

<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>
ML2254	ML2604	ML2611*	ML2611*	ML2709	ML2713
ML2746	ML2890	ML2891	ML2902	ML2918*	ML2918*

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Table 9: Composition of forty samples infected with HIV-1 subtype D with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)

<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>
ML1051*	ML1051*	ML1370*	ML1370*	ML1649	ML2155
ML2364	ML2366	ML2403	ML2405	ML2408	ML2449
ML2518*	ML2518*	ML2522	ML2569	ML2585*	ML2603
ML2633	ML2732	ML2756*	ML2756*	ML2767	ML2773
ML2773	ML2793*	ML2793*	ML2793*	ML2804	ML2814*
ML2818	ML2834	ML2834	ML2889*	ML2889*	ML2889*
ML2902	ML2912*	ML3029	ML2917*		

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Table 10: Composition of three samples infected with HIV-1 subtype G with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)

<u>Sample</u>	<u>Sample</u>	<u>Sample</u>
ML3038	ML3038*	ML3038*

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Table 11: Composition of eight samples infected with HIV-1 recombinants with non-significantly hypermutated HIV-1 provirus from partial pol sequencing (nucleotide positions 4150-4919)

<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>	<u>Sample</u>
ML1625	ML2185*	ML2185*	ML2264	ML2266	ML2332
ML2350	ML2797				

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

4.1.3 APOBEC3G/F-mediated hypermutation at HIV-1 vpu/env region is independent of the hypermutation at the HIV-1 pol region and *vice versa*

Subjects with available sample that demonstrated hypermutation at either vpu/env (117, 118) or pol region of the proviral HIV-1 genome were selected to further examine for hypermutation in the (as yet-unstudied) vpu/env or pol region of the HIV-1 genome. The majority of proviral DNA sequences further tested for APOBEC3G/F-mediated hypermutation were obtained from the same sampling time point as previous vpu/env or pol sequence, with the exception of subjects ML1053, ML1102, ML1857, and ML1957. In these subjects, the samples used to provide the proviral DNA were from a different time point. The list of subjects used in this study can be found in Table 12.

Of the twelve subjects for whom we had existing hypermutation data, nine subjects did not show APOBEC-mediated hypermutation at both sites of the HIV-1 genome (Table 13). This is an important finding as it suggests APOBEC-mediated hypermutation is independent at different regions of the HIV-1 genome. However, three subjects showed proviral hypermutation at both vpu/env and at pol regions of the HIV-1 genome (proviral DNA sequences of both regions examined matched date and time of sample collection). Proviral hypermutation observed in subject ML1578 had both detectable APOBEC3G ($p = 1.214e^{-5}$) and APOBEC3F ($p = 0.0014$) activity at the vpu/env site but only APOBEC3G activity ($p = 0.0370$) at the pol region. This suggests that APOBEC3G but not APOBEC3F activity can predominate at both vpu/env and pol regions in this particular subject. Another two subjects, ML256 and ML2801, had

Table 12: List of subjects where two regions of the HIV-1 genome was examined for APOBEC3G/F-mediated hypermutation

<u>Subject</u>	<u>Region of previous hypermutation detected</u>
ML256	pol
ML1053	vpu/env
ML1102	vpu/env
ML1578	vpu/env
ML1592	vpu/env
ML1847*	pol
ML1847*	pol
ML1857	vpu/env
ML1957	vpu/env
ML2503	pol
ML2503	pol
ML2645	pol
ML2766	pol
ML2801	pol
ML2814	pol
ML2818	pol
ML2818	pol
ML2862	pol
ML2871	pol
ML2885	pol
ML2891	pol
ML3033	pol

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Table 13: Comparison of hypermutated subjects at two different regions of the proviral HIV-1 genome

<u>Subject</u>	<u>ML date</u> (mm/dd/year)	<u>Evidence of hypermutation in Vpu/Env</u>	<u>Hypermutation activity Total^a/APOBEC3G^b/APOBEC3^b</u>	<u>Evidence of hypermutation in Pol</u>	<u>Hypermutation activity Total^a/APOBEC3G^b/APOBEC3^b</u>
ML256	10/26/2007	Y	0.1071/0.0064/0.7438	Y	0.0094/0.0014/0.2419
ML1053	05/03/1996	Y	0.0012/6.074e ⁻⁰⁵ /0.0644	N	0.9852/1/0.9214
ML1102	08/29/1995	Y	0.0029/1.466e ⁻⁰⁴ /0.1321	N	0.8363/0.8802/0.7581
ML1578	03/13/1996	Y	0.8124/1.214e-05/0.0014	Y	0.1051/0.0370/0.5711
ML1857	02/14/1995	Y	1.425e ⁻⁰⁴ /1.456e ⁻⁰⁴ /0.0038	N	0.8125/1/0.5998
ML1957	06/04/2002	Y	5.911e ⁻⁰⁹ /0.0251/6.44e ⁻¹⁰	N	0.9579/0.8187/0.9805
ML2503	06/25/2009	N	0.6187/0.8481/0.4845	Y	1.312e ⁻⁰⁹ /1.432e ⁻⁰⁹ /1.068e ⁻⁰⁵
ML2645	07/01/2009	N	0.5475/0.7400/0.4322	Y	0.0013/2.295e ⁻⁰⁶ /0.2800
ML2766	07/16/2007	N	1/1/1	Y	2.972e ⁻⁰⁷ /3.119e ⁻¹⁰ /0.0069
ML2801	06/19/2009	Y	0.0840/0.4436/0.0451	Y	1.327e ⁻⁰⁷ /1.791e ⁻⁰⁹ /0.0010
ML2871	06/09/2009	N	0.9795/1/0.8964	Y	0.0868/0.0436/0.0407
ML3033	07/15/2009	N	0.9086/0.8501/0.9133	Y	0.0632/0.0317/0.3543

^a General APOBEC-type hypermutation using Hypermut 2.0 analysis. Default settings of G to A substitutions with a downstream context of RD, where R is either A or G and D is A, T or G. Hypermutated sequences are defined as those with a p-value of <0.05.

^b Hypermut 2.0 analysis for specific APOBEC3G (settings of G to A substitutions with a downstream context of GD) and/or APOBEC3F (settings of G to A substitutions with a downstream context of AD) hypermutation. Hypermutated sequences are defined as those with a p-value of <0.05.

Significant values are highlighted in yellow. ML dates highlighted in grey indicate that vpu/env and pol sequences were from matched dates and time of blood collection, non-highlighted ML dates that were not closely matched differed between 1-3 years. Subject ML numbers highlighted in turquoise indicate these subjects had proviral hypermutation in both vpu/env and pol regions of HIV-1. Data italicized and in bold were obtained from Dr. Allison Land.

detectable APOBEC-mediated hypermutation at both the vpu/env and pol regions. APOBEC3G was the only hypermutation event detected at both proviral locations in ML256. The provirus of ML256 had APOBEC3G-mediated hypermutation at vpu/env sequences ($p = 0.0064$) and at pol sequences ($p = 0.0014$). The provirus of ML2801 had only detectable APOBEC3F-mediated hypermutation at vpu/env ($p = 0.0451$) and had hypermutation at pol sequences due to both APOBEC3G ($p = 1.791e^{-09}$) and APOBEC3F activity ($p = 0.0010$). APOBEC3G activity predominates at both regions of the HIV-1 genome in two of the three subjects (ML256 and ML1578) (Table 14). APOBEC3F activity is only occasionally present (as seen both in ML1578 and ML2801). Subject ML2801 showed activities of both APOBEC3G and APOBEC3F, but the differing p values suggest the activity of both proteins were not equally represented at both proviral locations and the activity of each molecule is independently controlled.

All other subjects in this group of twelve did not show statistically significant values of any type of hypermutation in both regions of HIV-1 proviral genome (pol or vpu/env). This suggests that APOBEC3G/F-mediated hypermutation occurring at one region is independent of hypermutation at another region of the HIV-1 genome.

Table 14: Three subjects that showed HIV-1 proviral APOBEC3G/F-mediated hypermutation at both vpu/env and at pol regions

<u>Subject</u>	<u>Hypermutation at vpu/env</u>		<u>Hypermutation at pol</u>	
	<u>APOBEC3G</u>	<u>APOBEC3F</u>	<u>APOBEC3G</u>	<u>APOBEC3F</u>
ML256	Yes	No	Yes	No
ML1578	Yes	Yes	Yes	No
ML2801	No	Yes	Yes	Yes

4.2 The presence of HIV-1 hypermutation at proviral *pol* gene correlates with higher CD4⁺ T cell counts in HIV-1⁺ Kenyan CSW

4.2.1 Hypermutation of proviral *pol* sequences introduces premature stop codons

To examine the functional effects of APOBEC-mediated hypermutation, we examined the coding differences caused by hypermutation. The sixteen proviral hypermutated sequences had numerous amino acid substitutions in the *pol* region of HIV-1. Often, nucleotide mutations occurred at tryptophan residues where G to A hypermutation caused the mRNA to change from a UGG codon to a stop codon (UAG, UAA, or UGA), disrupting the function of the gene. Table 15 lists the total number of premature stop codons found in the *pol* region for the sixteen samples with hypermutated provirus. Examples of the introduced amino acid and in-frame stop codon mutations are provided in Figure 17. These results suggest that premature stop codons mediated by APOBEC3G/F hypermutation could potentially lead to defective progeny virus. And that APOBEC3G/F mediates high levels of amino acid variants that could be defective in function.

Table 15: Number of premature stop codons present throughout the *pol* gene from sixteen samples with hypermutated pol proviruses

<u>Samples</u>	<u>Sequence length</u>	<u>No. of premature stop codons</u>
ML1847*	707	43
ML1847*	707	43
ML2503*	707	56
ML2503*	1765	139
ML256	707	46
ML2585	707	44
ML2645	1765	122
ML2766	707	54
ML2801	707	53
ML2814	707	43
ML2818	707	43
ML2818	707	42
ML2862	707	46
ML2871	707	45
ML2891	707	47
ML3033	707	44

*Sequences were obtained from multiple time points to observe if hypermutation activity changes over time.

Nucleotide	364					382					400					418					436					454		459				
Consensus	GGA	AGA	TGG	CCA	GTA	AAA	ACA	ATA	CAT	ACT	GAC	AAT	GGC	AGC	AAT	TTC	ACC	GGT	GCT	ACG	GTT	AGG	GCC	GCC	TGT	TGG	TGG	GCG	GGA	ATC	AAG	CAG
ML3033	A	...	R	...	GT	G	A	Y	...	G	A	A	AA	...	C	A
ML256	GTG	G	A	Y	...	G	A	R	A	...	C	A	
ML1847*	GT	R	A	C	...	G	A	A	...	C	A	
ML2503*	A	GT	G	A	A	C	...	G	A	A	A	...	C	A		
ML2585	GT	G	A	C	...	G	A	A	A	...	C	A		
ML2766	T	...	GT	G	A	C	...	G	A	A	A	...	C	A		
ML2801	GT	R	A	C	...	G	W	A	...	C	A	
ML2814	G	G	A	C	...	G	A	A	...	C	A	
ML2818*	R	G	A	C	...	G	A	A	...	C	A	
ML2862	R	R	A	Y	...	G	A	A	...	C	A	
ML2871	G	R	A	Y	...	G	A	A	...	C	A	
ML2891	C	...	GT	G	A	C	...	G	A	A	...	C	A	
ML2503*	A	GT	G	A	A	C	...	G	A	A	A	...	C	A		
ML2645	A	T	A	C	...	G	A	A	A	...	C	A		
Protein	122					128					134					140					146					152	153					
Consensus	G	R	W	P	V	K	T	I	H	T	D	N	G	S	N	F	T	G	A	T	V	R	A	A	C	W	W	A	G	I	K	Q
ML3033	.	.	W	.	?	.	V	V	?	?	.	.	.	G	A	A	?	K	.	.	.	?	W	.	?	.	.	.
ML256	V	V	?	?	A	A
ML1847*	V	V	?	?	A	A
ML2503*	.	.	W	.	.	.	V	V	?	.	.	.	S	A	A	W	W	
ML2585	V	V	?	A	A
ML2766	V	V	?	A	A	.	F
ML2801	V	V	?	A	A
ML2814	V	V	?	?	A	A	.	?
ML2818*	R	V	?	?	A	A	.	?
ML2862	R	?	?	?	A	A	.	?
ML2871	V	V	?	A	A
ML2891	V	V	?	A	A
ML2503*	V	V	?	?	A	A
ML2645	.	.	W	.	.	.	V	V	S	S	.	A	W	W	

Figure 17: Nucleotide and protein coding of sixteen samples with significantly hypermutated proviruses at HIV-1 pol. The upper panel illustrates the nucleotide positions 364 to 459 (chosen randomly) from each sample. The bottom panel illustrates protein residues (position 122 to 153) from the same sixteen samples. The top line at the upper and bottom panels illustrates the nucleotide consensus sequence and protein consensus, respectively. Identical sequence/protein residues are indicated with a dot (·); unknown residues are indicated with a question mark (?); while coding changes are shown with either a nucleotide or a one-letter amino acid code. Premature stop codons located in the protein residue are indicated with an asterisk (*). All G to A nucleotide substitutions in the upper consensus panel are highlighted in yellow, while all the stop codons in the bottom panel of the protein context are highlighted in yellow.

4.2.2 Proviral pol sequences correlate with CD4⁺ T cell counts

As APOBEC-mediated hypermutation may lead to non-functional proviruses, we investigated if there is a difference in disease progression as determined by CD4⁺ T cell counts in subjects with significant hypermutated and non-hypermutated proviruses. We examined the CD4⁺ T cell counts at the time the samples for pol sequences were taken and compared the mean CD4⁺ T cell counts between those who had hypermutation and those who did not. In Figure 18, CD4⁺ T cell counts for subjects (n=16) with hypermutated provirus were significantly higher than subjects with non-hypermutated provirus (n=219, p -value = 0.0388). This suggests that host factors such as APOBEC3 proteins may be playing an important role in protection against disease progression as overall hypermutated sequences correlated with a higher CD4⁺ T cell counts.

After observing that there is a clear relationship between CD4⁺ T cell counts and significantly hypermutated provirus in the pol region of HIV-1 and *vice versa*, we wanted to examine the entire sequence data set to determine if the relationship persists (Figure 19). The CD4⁺ T cell counts and adenine proportion (which can be used as an overall measure of APOBEC3 activity) were correlated in an unbiased case-wise comparison for subjects with available CD4⁺ T cell counts (n = 235). The proviral adenine proportion and CD4⁺ T cell counts correlated strongly ($p < 0.0001$), but with a weak r -value of 0.05.

This set of data shows that there is a significant association between disease progression (in terms of CD4⁺ T cell counts) between two sets of groups - subjects with hypermutated proviruses in pol and those with non-hypermutated proviruses. This

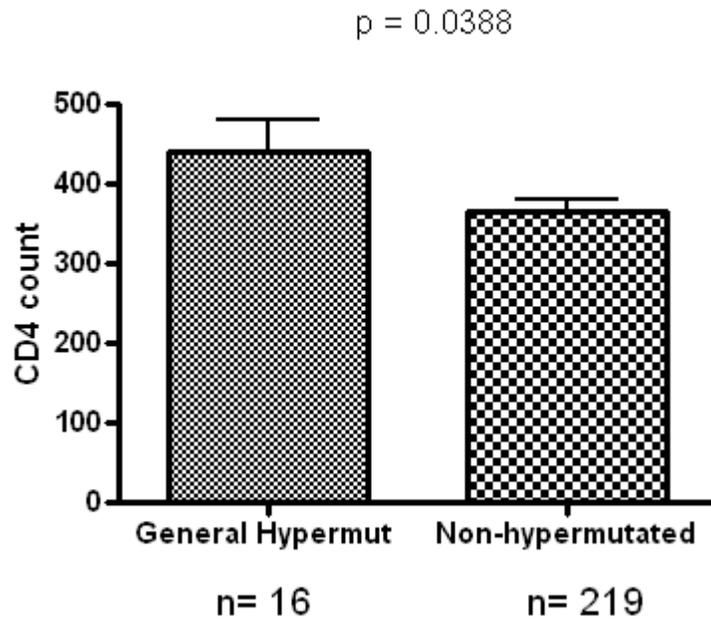


Figure 18: Association of higher CD4⁺ count with hypermutated proviral DNA sequences in pol. CD4⁺ counts are compared between subjects with hypermutated proviral DNA sequences and subjects without hypermutated proviral DNA sequences. Groups are compared by a Mann-Whitney test ($p = 0.0388$). The height of the bar indicates the mean, and the error bar represents the standard error of the mean.

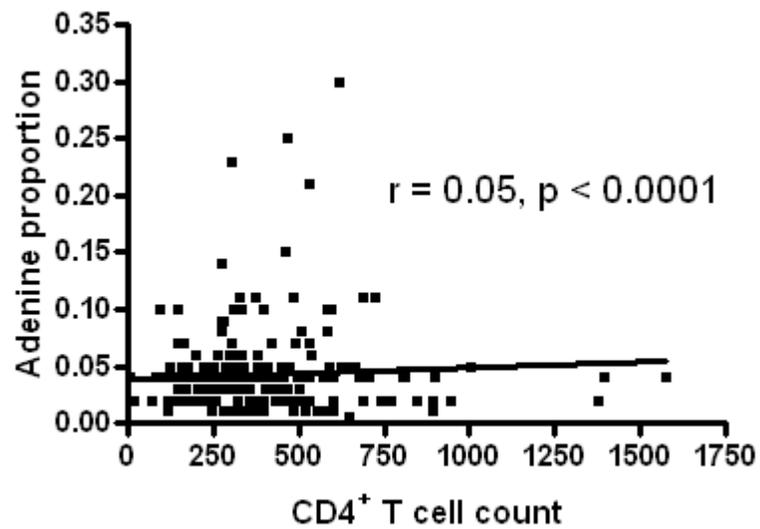


Figure 19: Correlation of adenine proportion in the pol region of proviral HIV-1 with CD4⁺ T cell counts from 235 subjects. Each data point represents one subject. The adenine proportion and CD4⁺ T cell counts were correlated strongly, but weakly with a correlation coefficient, $r = 0.05$ and a p -value < 0.001 .

confirms previous data from women in the Pumwani CSW cohort (118), suggesting that APOBEC3 activity in women from this cohort may play a protective role in disease progression.

All sixteen subjects with HIV-1 proviral hypermutation in pol showed statistically significant values of APOBEC-mediated hypermutation (Table 6). This results in a disruption in pol due to the nucleotide G to A substitutions. Amino acid mutations and premature stop codons are introduced into the coding region of mRNA leading to defective provirus. When these subjects were examined in relation to HIV-1 disease progression, they showed higher CD4⁺ T cell counts than subjects with non-hypermutated proviruses. This strongly suggests that HIV-1 disease progression is related to APOBEC-mediated proviral hypermutation.

4.3 HIV-1⁺ women with hypermutation at the proviral HIV-1 genome (env, pol, and vpu) have increased APOBEC3G expression

4.3.1 Rationale

A subset of women from the Pumwani CSW cohort demonstrated hypermutation in vpu/env or pol of the proviral HIV-1 genome in PBMCs. Theoretically; this could be due to higher endogenous APOBEC3G protein expression. If these subjects have increased APOBEC3G protein expression and overwhelm the activity of Vif (which counteracts APOBEC3G); this could lead to higher levels of hypermutation in these subjects. The purpose of this experiment was to observe, *ex vivo*; the expression levels of endogenous APOBEC3G from subjects with hypermutated proviral sequences compared to those with non-hypermutated sequences. A total of twenty subjects who were demonstrated to have hypermutated (n = 12) and non-hypermutated (n = 8) proviruses were analyzed for endogenous APOBEC3G and Vif protein expression from PBMCs by Western blot analysis (Table 16).

Table 16: PBMC list of subjects

<u>Subjects</u>	<u>Infection subtype</u>	<u>Significantly hypermutated at pol</u>	<u>Significantly hypermutated at vpu/env</u>	<u>Region of hypermutation examined</u>
ML1102	D	Yes	Not determined	vpu/env
ML1230	A1	Yes	Not determined	vpu/env
ML1592	A1	Yes	Not determined	vpu/env
ML1847	A1	Not determined	Yes	pol
ML1857	C	Yes	Not determined	vpu/env
ML1970	A1	Yes	Not determined	vpu/env
ML1975	A1	Yes	Not determined	vpu/env
ML2019	C	Yes	Not determined	vpu/env
ML2503	A1	Not determined	Yes	pol
ML2645	A1	Not determined	Yes	pol
ML2713	C	No	--	--
ML2715	A1	No	--	--
ML2761	A1	No	--	--
ML2763	A1	No	--	--
ML2766	A1	Not determined	Yes	pol
ML2767	D	No	--	--
ML2768	A1	No	--	--
ML2770	A1	No	--	--
ML2779	A1	No	--	--
ML2818	A1	Not determined	Yes	pol

4.3.2 *Ex vivo* expression of HIV-1 Vif and endogenous APOBEC3G proteins in PBMCs of HIV-1⁺ women with APOBEC-mediated vpu/env/pol hypermutation

To test if subjects with proviral hypermutated DNA sequences had increased APOBEC3G and/or decreased Vif expression, protein was isolated from PBMCs from subjects in Table 16 and examined by Western blot (Figure 20). A set of controls were first established to illustrate that the antibodies, vif2222-vif809 and ApoC17, were able to specifically bind Vif and APOBEC3G proteins, respectively. Our controls were set up in a transfection system as follows: untransfected HEK 293 T cells, HEK 293 T cells transfected with the expression vector pcDNA-HVif (expressed only wt Vif protein), HEK 293 T cells transfected with the expression vector pCMV-HA (expressed only wt APOBEC3G protein), and HEK 293 T cells transfected with both expression vectors pcDNA-HVif (expressed wt Vif protein) and pCMV-HA (expressed wt APOBEC3G protein), (Lanes 1-4). Based on Figure 20, untransfected HEK 293 T cells did not show any protein expression at either 46 kDa or 23 kDa the molecular weight of APOBEC3G and Vif, respectively (Lane 4). Transfected whole cell lysates of wt Vif and wt APOBEC3G alone illustrated appropriate bands at the 23kDa and 46 kDa markers, respectively (Lanes 2 and 3). Finally, HEK 293 T cells that were co-transfected with pCMV-HA and the pcDNA-HVif expression vectors showed two different products (Lane 1) – with a strong band seen at 23 kDa marker and a weaker APOBEC3G expression at 46 kDa. Based on the above results, the wt controls show strong band intensities illustrating the specificity of antibodies, vif2222-vif809 and ApoC17. HEK 293 T cells do not express endogenous APOBEC3G hence; it is why we do not observe any expression of proteins when antibodies are immunoblotted against Vif or

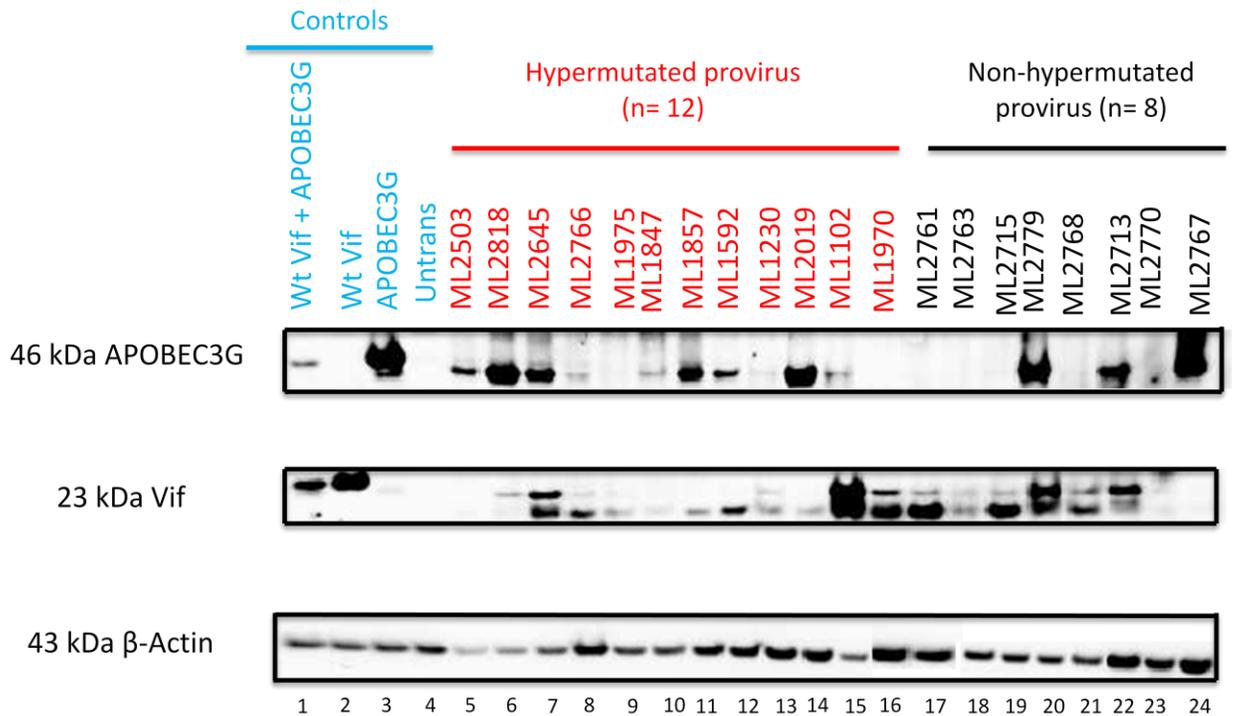


Figure 20: APOBEC3G and Vif expression in PBMC cell lysates. Frozen PBMCs were lysed with RIPA lysis buffer. Whole cell lysates were collected to detect Vif and APOBEC3G expression. All samples were subjected to reducing on a 4-12% SDS-PAGE gradient gel and immunoblotted for detection of Vif using primary antibodies anti-vif809 (1:1000) and anti-vif2221 (1:000) in a 1:1 ratio. Membrane was also immunoblotted for APOBEC3G using anti-ApoC17 (1:14000) and β -actin (1:7500). Controls (labelled in blue) shown in lanes 1-4 illustrating appropriate antibodies are able to detect proteins of interest. Subjects with hypermutated proviruses labelled in red and non-hypermutated proviruses labelled in black as indicated on top of blot. Lanes labelled from 1-24 as indicated at bottom of blot. This is representative of one independent experiment.

APOBEC3G proteins. The cytoskeleton protein, actin was used as a comparison to confirm that each sample had equal amounts of protein loaded into wells. This experiment could only be conducted once as obtaining PBMCs from subjects in the laboratory were difficult to obtain due to lack of sample availability.

Ex vivo, we measured endogenous APOBEC3G and endogenous Vif proteins in PBMCs from infected subjects to illustrate the protein expression differences in subjects with hypermutated (n=12) and non-hypermutated (n=8) proviruses. Detectable APOBEC3G expression was observed in the majority of samples and the level of endogenous APOBEC3G varied strongly between subjects with and without hypermutation at the HIV-1 genome. Some subjects had no detectable levels of endogenous APOBEC3G. In subjects with hypermutated proviruses, the majority of subjects (10/12) had detectable endogenous APOBEC3G (Lanes 5-12, 14, and 15) while in a few subjects (2/12) endogenous APOBEC3G was undetectable (Lanes 13 and 16). In subjects with non-hypermutated proviruses, endogenous APOBEC3G was undetectable for the majority of subjects (5/8) (Lanes 17-19, 21 and 23). There are a few subjects such as ML2779, ML2713, and ML2767 (Lanes 20, 22, and 24) who were observed to have detectable APOBEC3G expression (3/8). To determine whether there was a significant difference between expected and observed frequencies of our two groups of subjects (those with hypermutated provirus versus those with non-hypermutated provirus at either vpu/env or pol) we conducted an analysis by chi-square test. Based on the above results, we found that there was a significant difference observed between the two groups with a *p*-value of 0.03 with a chi-square value of 4.432 and one degree of freedom.

Strong band intensities at 46 kDa were detected in 33% in subjects with hypermutated proviruses. These subjects had the highest level of endogenous APOBEC3G expression (Lanes 6, 7, 11, and 14). There was moderate endogenous APOBEC3G expression in 17% of subjects with hypermutated provirus (Lanes 5 and 12), while the rest had low levels of endogenous APOBEC3G (Lanes 8, 10, 13, and 15). In subjects with non-hypermutated proviruses, 38% had the high level of endogenous APOBEC3G expression, similar to subjects with hypermutated proviruses (Lanes 20, 22, and 24). One subject (13%) demonstrated very low levels of APOBEC3G (Lane 23). The remainder of subjects (50%) had non-detectable levels of endogenous APOBEC3G (Lanes 17-19, and 21). This data suggests that subjects with non-hypermutated provirus may not express detectable levels of endogenous APOBEC3G proteins or it may be present at extremely low levels that cannot be detected by Western blot.

Next, we determined the levels of Vif expression by infected, hypermutated and non-hypermutated PBMC samples. The Western blot showed that endogenous Vif proteins were expressed in most subjects from both groups at varying levels (Table 17, fourth column). By visualizing the data, endogenous Vif seems to have higher expression in subjects with non-hypermutated proviruses when compared to subjects with hypermutated proviruses. To determine if this is the case, we conducted a chi-square analysis and found no association between Vif expression and hypermutated provirus ($p = 0.4190$). The large variation in Vif expression made it difficult to establish a relationship between Vif expression and functional interactions with APOBEC3G. Thus, trying to understand the ability of endogenous Vif to degrade APOBEC3G and draw conclusions

Table 17: The effect of endogenous Vif protein expression of subjects on degradation of endogenous APOBEC3G protein

	<u>Subject</u>	<u>Strength of APOBEC3G expression*</u>	<u>Strength of Vif expression*</u>	<u>Detectable ability of Vif to degrade APOBEC3G</u>
Hypermuted provirus	ML2503	(+)	(-)	No
	ML2818	(+++)	(-)	No
	ML2645	(+++)	(+)	No
	ML2766	(+)	(++)	Yes
	ML1975	(+)	(+)	Yes
	ML1847	(+)	(+)	Yes
	ML1857	(+++)	(+)	No
	ML1592	(++)	(++)	No
	ML1230	(-)	(+)	Yes
	ML2019	(+++)	(+)	No
	ML1102	(+)	(+++)	Some
	ML1970	(-)	(+++)	Yes
Non-hypermuted provirus	ML2761	(-)	(+++)	Yes
	ML2763	(-)	(+)	Yes
	ML2715	(-)	(+++)	Yes
	ML2779	(+++)	(++)	No
	ML2768	(-)	(+)	Yes
	ML2713	(++)	(+)	No
	ML2770	(-)	(-)	Maybe
	ML2767	(+++)	(-)	No

*The strength of expression from subjects with Vif and/or wt APOBEC3G protein is denoted by a negative sign (-) or a positive sign (+). The (-) sign represents low and/or weak expression of protein. The (+), (++) and (+++) signs represents low to moderate to highest protein expression.

about its expression in the hypermutated proviral and non-hypermutated proviral group is problematic. Table 17 illustrates the protein expression patterns of both endogenous proteins. Theoretically, we would expect subjects with hypermutated provirus to have low or dysfunctional Vif expression and an impaired ability to degrade endogenous APOBEC3G and we should observe more endogenous APOBEC3G expression. However, our data indicates that 50% of subjects in both groups reflect this situation. It is difficult to determine a relationship between Vif and APOBEC3G due to the variability existing between the virus (Vif) expression or the host APOBEC3G expression.

4.4 *In vitro* expression of HIV-1 Vif in subjects with vpu/env hypermutation was reduced compared to non-hypermutated subjects

4.4.1 Rationale

A subset of women in the Pumwani CSW cohort were found to have statistically significant (p value ≤ 0.05) high levels of APOBEC-mediated hypermutation at the vpu/env region of proviral HIV-1 genome. While the majority of HIV⁺ women of the same cohort exhibited relatively low levels of APOBEC-mediated hypermutation. These observations led us to speculate that HIV-1 Vif proteins present in subjects with significantly high level of hypermutation may be deficient in its ability to fully restrict APOBEC3G function. When, vif sequences from subjects with hypermutated and non-hypermutated proviruses were sequenced (by Dr. Allison Land) there did not appear to be any obvious differences in sequences in regions thought to be important in Vif function (Figure 21). However, any sequence variation may cause functional differences in the ability of Vif proteins to bind APOBEC3G proteins. Whether the Vif protein expression

or function differs between the two groups of subjects was tested in an *in vitro* model system.

A total of six vif constructs were available (vif sequences cloned into SCMV-IN expression plasmids). Three vif constructs were cloned from subjects that had significantly high levels of hypermutation in the *vpu/env* genes; and another three vif constructs were cloned from HIV⁺ subjects that did not have hypermutation in the *vpu/env* genes. All six subjects were shown to lack significant APOBEC-mediated hypermutation at the vif region of HIV-1 genome (Table 18).

4.4.2 Verification of plasmid construct by restriction enzyme digestion

The correct size of vif and APOBEC3G inserts in the purified plasmids were verified using restriction digests with the two restriction enzymes *pst* I and *xba* I (Figure 22). All plasmids contained inserts of the expected size. The inserts were sequenced by Dr. Allison Land.

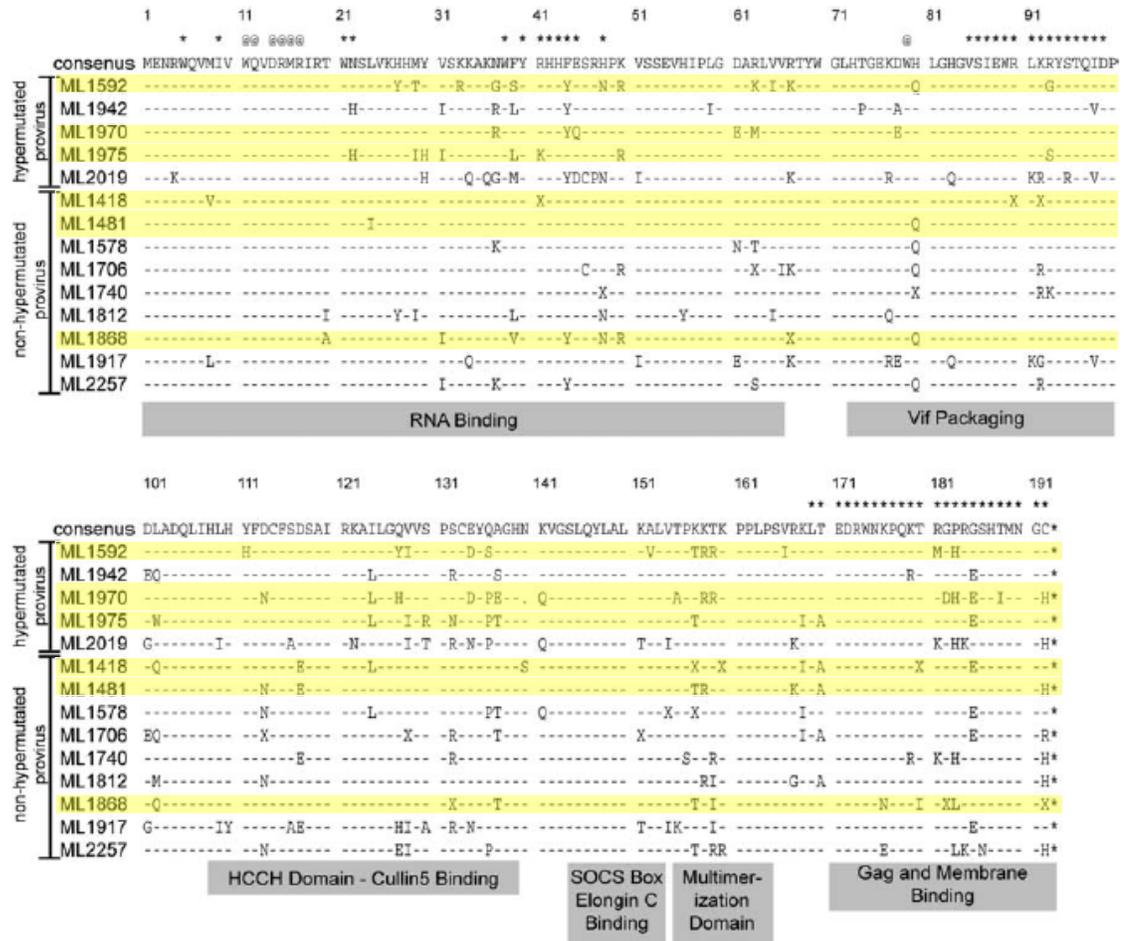


Figure 21: Comparison of *vif* sequences with subjects that have hypermutated and non-hypermutated HIV-1 proviral vpu/env sequences (117). The consensus sequence for the upper and bottom panel can be found on the top line of each panel; matched sequence residues are indicated with (-), one-letter amino acid code denote non-identical residues. Multiple DNA nucleotides in the corresponding codon are denoted with an “X” indicating ambiguous residues. Hypermutated sequences (ML1592, ML1970, and ML1975) and non-hypermutated sequences (ML1418, ML1481, and ML1868) were cloned into the vector, SCMV-IN, by Dr. Allison Land. Plasmids were kindly donated by Dr. Land. Sequences from subjects used in this study are highlighted in yellow. Figure was reproduced (slight modifications made to original) with permission from Dr. Allison Land on January 21, 2011.

Table 18: Patient list and proviral HIV-1 hypermutation characteristics at vpu/env and vif

Subject^a	Significant hypermutation in vpu/env^a	Vpu/env hypermut significance^a	Significant hypermutation in vif^a	Vif hypermut significance^a
ML1418	No	0.2984	No	0.511
ML1481	No	0.6066	No	0.798
ML1592	Yes	2.660×10^{-10}	No	0.974
ML1868	No	0.3946	No	0.979
ML1970	Yes	0.0058	No	0.971
ML1975	Yes	8.894×10^{-5}	No	1

^aTable adapted from Dr. Allison M. Land (117). Table was modified from the thesis titled: "The role of HIV-1 recombination and APOBEC3F/G-mediated hypermutation in HIV-1 pathogenesis." Permission obtained from Dr. Allison Land on January 21, 2011.

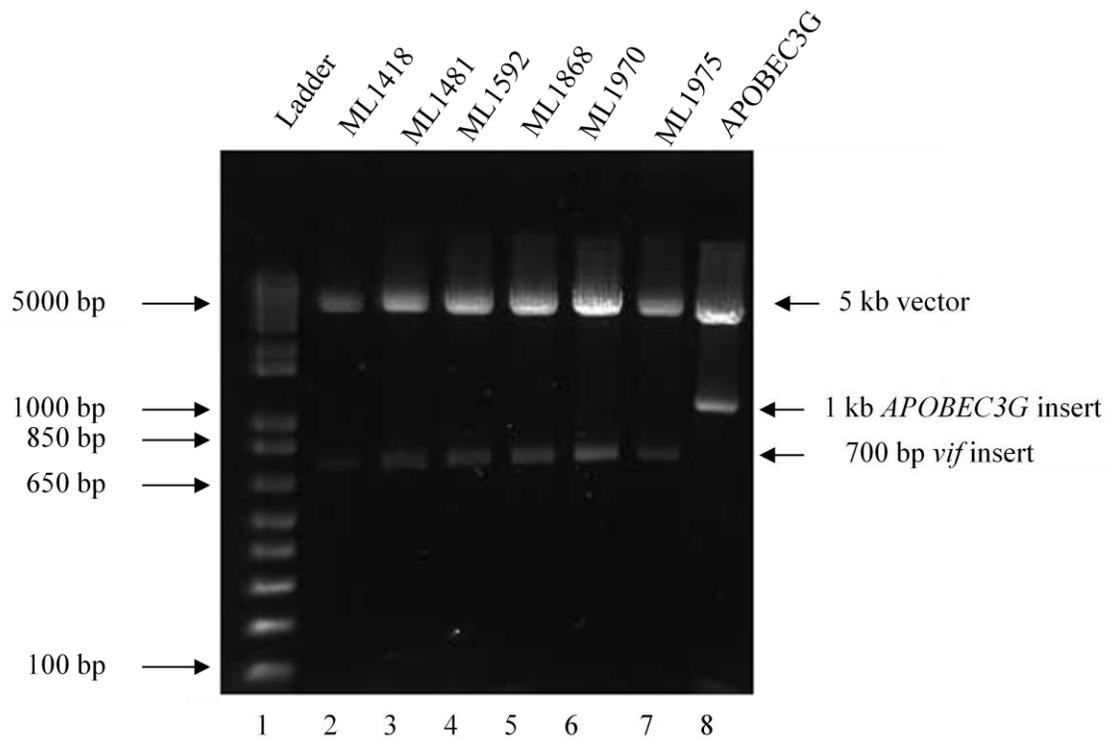


Figure 22: 1% agarose gel of *pst* I and *xba* I digest of expression vectors containing the *vif* and APOBEC3G insert. Expression vectors demonstrate that all inserts were of correct length and present in vector. Lane 1, DNA ladder; lanes 2-7 *vif* inserts shown at 700 bp and SCMV-IN vector without *vif* insert (5 kb) from different subjects; lane 8 APOBEC3G shown at 1 kb and pCMV-IN without APOBEC3G insert.

4.4.3 Co-expression of wt APOBEC3G with HIV-1 Vif, isolated from subjects with HIV-1 vpu/env proviral hypermutation-mediated by APOBEC: vif isolated from subjects with or without hypermutation had similar effects on wt APOBEC3G expression

A set of controls were first established to confirm that our transfection system was in working order and it was tested by Western blot (Figure 23). As described in section 4.3.2, the plasmids, pcDNA-HVif and pCMV-HA, were used as controls and were (co-) transfected into HEK 293 T cells to illustrate that the antibodies specifically bind to wt Vif and wt APOBEC3G proteins. Untransfected HEK 293 T cells did not show any protein expression at either the 46 kDa or 23 kDa mark (Lane 4). HEK 293 T cells do not express endogenous APOBEC3G. The structural protein glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as an index comparison to confirm that equal amount of protein was loaded into the wells. This experiment was done in triplicates and each experiment produced the same results in all three occasions.

All six HIV-1 Vif constructs were transfected into HEK 293 T cells (2 ug plasmid DNA). All six HIV-1 Vif constructs showed reduced level of expression (Lanes 5, 7, 9, 11, 13, and 15) compared to the wt Vif construct (pcDNA-HVif) (Lane 1). Detectable Vif expression was observed except for ML1418 (Lane 5) and ML1868 (Lane 11) where the expression level of Vif constructs, was barely detectable. Thus it seemed to appear that there was higher Vif expression in subjects with hypermutated provirus than non-hypermutated provirus at vpu/env of HIV-1 genome. Expression of the vif constructs from ML1592, ML1970, and ML1975 was slightly higher than ML1418, ML1481, and ML1868 non-hypermutated provirus. As previously mentioned,

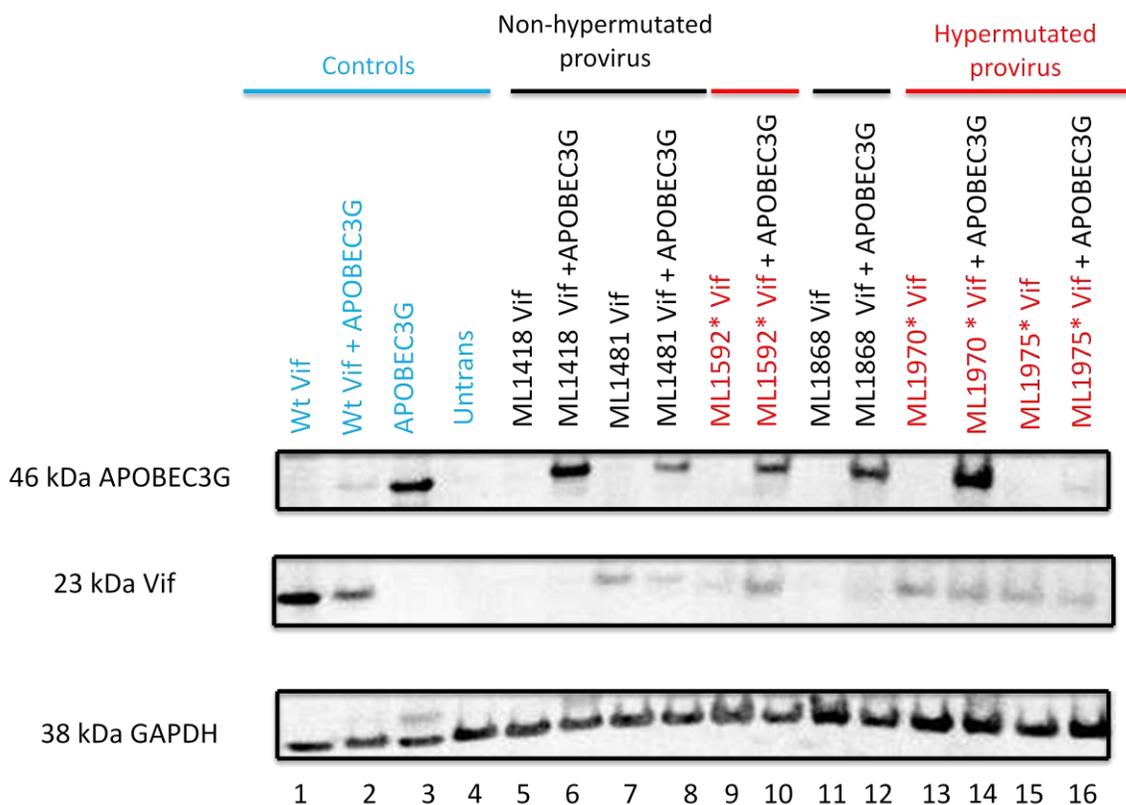


Figure 23: Expression of wt Vif, subject Vifs, and APOBEC3G proteins. HEK 293 T cells were transfected with wt vif, wt APOBEC3G, subject vifs alone, and co-transfected with wt APOBEC3G and appropriate subject Vif. Whole cell lysates were collected to detect Vif and APOBEC3G proteins. All samples were subjected to reducing on a 4-12% SDS-PAGE gradient gel and immunoblotted for detection of Vif using primary antibodies anti-vif809 (1:1000) and anti-vif2221 (1:000) in a 1:1 ratio. Membrane was also immunoblotted for APOBEC3G using anti-ApoC17 (1:14000) and anti-GAPDH (1:1000). Samples as indicated on top of blot and lane numbers indicated on bottom of blot. Samples marked with asterisk (*) and labelled in red represent subjects with hypermutated provirus at vpu/env (ML1592, ML1970, and ML1975). Samples labelled in black represent subjects with non-hypermutated provirus at vpu/env (ML1418, ML1481, and ML1868). Samples in blue represent controls for this study. This is representative of three independent experiments.

HEK 293 T cells do not express APOBEC3G and hence, when (co-)/transfected with the expression plasmid pCMV-HA (1 ug plasmid DNA), wt APOBEC3G was expressed at high levels.

When the effect of wt Vif on wt APOBEC3G expression was observed in combination with wt Vif and compared to the expression of wt APOBEC3G alone, it was obvious that the expression of wt Vif substantially reduced the expression of wt APOBEC3G (Lanes 2 vs. Lane 3). Vif from subjects with non-hypermutated vpu/env provirus (ML1418 and ML1868) had no effect on wt APOBEC3G expression (Lanes 6 and 12), while APOBEC3G expression in ML1481 was reduced a significant amount (Lane 8). Vif from hypermutated provirus in subject ML1592 knocked down APOBEC3G expression moderately (Lane 10), while ML1970 did not reduce APOBEC3G expression at all (Lane 14), and the APOBEC3G from ML1975 was knocked down substantially from this patient. Here we can conclude that the ability of Vif to degrade APOBEC3G varied independent of test group (summarized in Table 19).

Four of the six Vif variants that were cloned into SCMV-IN expression vectors gave similar expression levels of Vif protein when compared to one another (with the exception of ML1418 and ML1868). Even though the expression of Vif protein was reduced in each subject versus the wt control, there was an apparent ability of Vif from some subjects in both groups to inhibit or degrade APOBEC3G expression. However this was inconsistent and could be affected by differences in Vif expression, APOBEC3G

Table 19: The effect of Vif protein expression of subjects on degradation of wt APOBEC3G protein

	<u>Subject</u>	<u>Strength of Vif expression*</u>	<u>Strength of APOBEC3G expression*</u>	<u>Detectable ability of Vif to degrade APOBEC3G</u>
Vif from non-hypermutated vpu/env HIV-1 provirus	ML1418	(-)	(+++)	No
	ML1481	(++)	(+)	Yes
	ML1868	(-)	(+++)	No
Vif from hypermutated vpu/env HIV-1 provirus	ML1592	(+)	(++)	Maybe
	ML1970	(++)	(+++)	No
	ML1975	(++)	(-)	Yes

*The strength of expression from subjects with Vif and/or wt APOBEC3G protein is denoted by a negative sign (-) or a positive sign (+). The (-) sign represents low and/or weak expression of protein. The (+), (++) and (+++) signs represent low to moderate to highest protein expression.

expression or both. When comparing the protein co-expression of Vif and wt APOBEC3G in each subject to one another, there seems to be some striking variability in wt APOBEC3G expression between the subjects making it difficult to detect any effect of Vif from subjects with hypermutated or non-hypermutated provirus on APOBEC3G expression.

5.0 Discussion

The incorporation of the host HIV-1 restriction factors, APOBEC3G and APOBEC3F, into HIV-1 virions leads to the editing of viral genomic DNA. These nucleotide substitutions alter the amino acid sequences and sometimes lead to premature stop codons in the resulting protein. The effects of cytidine deamination of viral genomic DNA maybe critically important in understanding HIV-1 disease progression, but this is still heavily debated in the literature. Not only may cytidine deamination play a significant role in disease progression but the interaction between HIV-1 Vif and APOBEC3G/F proteins are essential in understanding the role of HIV-1 pathogenesis. However, the mechanism of interaction between HIV-1 Vif and APOBEC3 proteins are still unclear and even though there are several reports on the antiviral roles of APOBEC3 proteins on HIV-1, the mechanisms between host and virus interaction need to be studied more extensively to understand what is happening at the molecular as well as the host level.

5.1 Sequencing data suggest that hypermutation at HIV-1 proviral pol region is mediated by APOBEC3G/F: the significance

By examining hypermutation in sequences of pol from the Pumwani CSW cohort, we were able to show that sixteen out of two hundred forty-one subjects or 7% of all sequences demonstrated statistically significant effects of cytidine deaminase activity. Previous literature show that overall 5-12% patient sequences exhibit hypermutation (146, 151, 195). A study conducted by Piantadosi *et al.* compared single-copy gag sequences based on viral subtypes and found a total of 9.2% of all sequences observed were hypermutated (43/465 sequences). Hypermutation appears to vary from patient to patient as a study by Gandhi *et al.* observed that individuals on HAART treatment have a greater cytidine deaminase activity (hypermutation) of 4-48% in gag (64), while the same study indicated that one elite suppressor (individuals who are able to maintain an undetectable viral load) had more than 80% of its genome hypermutated (64). Vázquez-Pérez *et al.* observed individuals who had low viral loads when compared to HIV⁺ individuals and their sequences were hypermutated 8.86% and 7.9%, respectively (196). This suggests that individuals with low viral loads have increased rates of hypermutation. Furthermore, this study showed that increased APOBEC3G mRNA expression was positively correlated with G to A nucleotide substitution ($Rho = 0.43$, $p = 0.0226$) (196). This suggests that hypermutation may play a role in controlling viremia and thus disease progression.

Similar to our findings, others have also observed hypermutation in other proviral sequences of HIV-1 gag, env, nef, vif, LTR (39, 64, 106, 151, 191) and only a few studies have conducted analysis on near full-length genomes (9, 107, 189). Ulenga *et al.*

conducted proviral hypermutation studies on three regions of the HIV-1 genome (gag, env, and vif) from twenty-nine patients and found varying levels of G to A hypermutations at each site (18%, 15%, and 22%, respectively) (191). Land *et al.* conducted proviral hypermutation studies in HIV-1 vpu/env region from women in CSW cohort and found 5% of patients (13/240) had hypermutation in this region (118), while this thesis showed that 7% of all patients (16/241) from the same cohort had proviral hypermutation in the last 705 bases of pol. These three studies are in agreement with the study conducted by Suspene *et al.* that hypermutation differs in specific HIV-1 genes and supports the idea that hypermutation occurs at different rates along the HIV-1 genome (189).

This thesis has only focussed on partial pol sequencing; it would be useful to conduct sequencing on the full length of HIV-1 pol, as this may identify more subjects with proviral hypermutation. This would increase the number of subjects with full-length proviral pol hypermutation as the majority of the gene is found in the upper portion of the twin gradients (demonstrating very active APOBEC activity) as shown by Suspene *et al.* in Figure 9 (189). Kijak *et al.* conducted full-length proviral hypermutation studies on twenty-four patients and found that the level of hypermutation along the HIV-1 genome varied among patients who had hypermutation (75% or 18/24 patients had hypermutation across the genome) (107). Kijak *et al.* observed these eighteen patients were in agreement with Suspene *et al.* study that twin peaks were observed in hypermutated proviral sequences, not exactly at the same nucleotide positions but relatively close (between nucleotide position 3000-4000 and again at nucleotide position 7000-9000) (107). Unfortunately, this group was unable to state the percentage of G to A hypermutation in

each region of HIV-1; however they noted that hypermutation was lowest in HIV-1 gag and the hypermutation steadily increased till the RNaseH domain in HIV-1 pol (107). This suggests that examining full-length genome sequences may be a better option to evaluate for APOBEC-mediated hypermutation as this would better identify genome-wide hypermutation rates. Conducting full-length hypermutation studies across the genome would be more favourable rather than the examination of one or few regions of the HIV-1 genome as it would benefit a better evaluation of the relationship between hypermutation and HIV-1 disease progression.

The results shown in Table 6 suggest that the majority of subjects with hypermutated provirus had overall evidence of general hypermutation when using the Hypermut 2.0 tool. We have found that some subjects do demonstrate statistically significant evidence of hypermutation overall and APOBEC3G or APOBEC3F hypermutation, but rarely both. The importance of this does not mean that both enzymes would act equally on a virus. For example, a subject may express higher levels of one protein than the other, or the virus may encode a Vif protein that is able to better counteract one APOBEC3 protein over the others. Also, there are instances where overall (general) hypermut significance is valid, but neither APOBEC3G nor APOBEC3F is significant. This may be due to the possibility that the program counts more events in terms of its dinucleotide context (GG and GA contexts). When enzyme activity is split between the two proteins, significance is lost. The general hypermut value calculates all known G to A changes in the particular region that is being examined (GG to GA, GA to AA, GC to AC, and GT to AT), hence this includes APOBEC3 protein activity (GG to GA, GA to AA) and mutations that are not APOBEC3 related which are caused by the

dinucleotide contexts of GC to AC, and GT to AT (112). In turn, if only general hypermut significance warrants over APOBEC3G and APOBEC3F then adenine substitutions are likely due to other factors than APOBEC3 proteins.

5.2 The presence of hypermutation at HIV-1 proviral pol region correlates with higher CD4⁺ T cell counts: the significance in disease progression

APOBEC3G and APOBEC3F activity may play a role in affecting HIV-1 disease progression (118, 146), and they have also been hypothesized to play role in resistance to infection due to its polymorphic variations (143). Previous data from our group showed that a subset of HIV-1⁺ women (n=13) with detectable hypermutation activity in vpu/env had significantly higher CD4⁺ T cell counts ($p = 0.0052$) than other subjects who did not have hypermutation (n=184) in the vpu/env region (118). Our group further investigated the entire data set of subjects to see if adenine proportion (an overall measure of APOBEC activity) correlated with CD4⁺ T cell counts and found a significant correlation ($r = 0.1411$, $p = 0.042$) (118). It has been established that CD4⁺ T cell counts can be a good marker in predicting disease progression of HIV⁺ individuals. Numerous studies have used CD4⁺ T cell counts as a reliable method in describing disease progression when comparing hypermutated sequences (97, 118, 144, 151, 162). The CD4⁺ T cell counts between subjects with hypermutated and non-hypermutated proviral DNA sequences in HIV-1 pol were elevated in those with increased hypermutated provirus ($p = 0.0388$). This suggests that host factors such as APOBEC3G-mediated hypermutation may be playing an important role in protection, as subjects with overall hypermutated proviral sequences have a higher CD4⁺ T cell count and higher CD4⁺ T cell counts suggests a

healthier status.

In addition to CD4⁺ T cell counts as an indicator for predicting disease progression, other studies have used a second marker, viral load, to measure the level of viremia in plasma samples of subjects. The methods used to examine a relationship between disease progression and APOBEC activity include: measuring mRNA levels of APOBEC3G by RT-PCR (144, 162), examining protein expression levels, and calculating the evidence of APOBEC-mediated hypermutation (the last two markers were examined in this thesis). Jin *et al.* were the first group to correlate APOBEC3G mRNA levels to both viral load data and to CD4⁺ T cell counts (97). They were able to show that APOBEC3G mRNA levels were positively correlated to CD4⁺ T cell counts and that an inverse relationship existed between APOBEC3G mRNA levels and viral load (97). Unfortunately, our laboratory does not routinely capture viral load data due to cost considerations. A study by Land *et al.* showed that there were no correlations between viral loads of hypermutated and non-hypermutated subjects (but viral load data was not conducted on fresh samples), while another study by Ulenga *et al.* published in the same year showed no relationship between viral loads and G to A hypermutations (118, 191). Based on the above studies and many other published literature, there seems to be some controversy in developing a linkage between HIV-1 proviral hypermutation and disease progression. However, data from this thesis provides confirmatory data that HIV-1 proviral APOBEC-mediated hypermutation does associate with CD4⁺ T cell counts.

The majority of studies examining proviral hypermutation tend to look at only two or three of the variables (APOBEC, CD4 T cell counts, and viral load). There still seems to be conflicting data between different studies and therefore the scientific community

cannot come to a clear conclusion about APOBEC3G/F in playing a role in HIV-1 disease progression. This may be due to patient-to-patient variability and leans to the confounding possibility that numerous other host factors are playing a role in disease progression. If we were able to account for both viral load and CD4⁺ T cell data in our study, we could come to a more specific conclusion regarding HIV-1 disease progression in hypermutated subjects. Further studies for this set of subjects should determine the level of circulating virus as this may be an important factor in assessing a correlation between hypermutation and non-hypermutated subjects. Finally, it would also be interesting to see what the APOBEC3G mRNA levels are in these subjects. Analyzing all three variables will give a stronger and more conclusive answer in the role of hypermutation in disease progression.

5.3 The presence of hypermutation at HIV-1 proviral vpu/env region is independent of that at the pol region: possible mechanisms and significance

Previous findings from our lab found hypermutation in vpu/env region from proviral HIV-1 DNA. We determined if individuals demonstrate significant hypermutation at both the vpu/env and the pol region. Based on sample availability, some subjects were sequenced twice at the two different regions - the pol and vpu/env and we found that there did not appear to be significant shared hypermutation patterns between vpu/env/pol and *vice versa*. Only three subjects out of twelve did show significant proviral hypermutations at both vpu/env and pol regions of the HIV-1 genome. The entire vpu/env region was sequenced in all subjects of this study, however only partial pol sequencing (the last 705 bp of the gene) was conducted in these subjects. Subject

ML2503 was one of two subjects that underwent near full length pol sequencing (out of a total of twenty-eight subjects) that had proviral hypermutation. This suggests that if we were able to get longer sequence data from HIV-1 pol, we may see a better correlation between hypermutation at multiple sites. Due to difficulties such as time constraints, length of gene, and success rate associated with near full-length sequencing of pol (2 kb in length), we focussed on partial sequencing of pol (705 bp in length). It is unfortunate that few subjects were available to conduct near full length pol sequencing at the time of study and we were unable to obtain samples from the same subjects at multiple time points to use for partial pol sequencing to see if hypermutation persists. However this particular subject, ML2503, did show hypermutation from earlier partial pol sequencing. This suggests that full length sequencing may be the better option to properly examine APOBEC-mediated hypermutation.

It is possible that the rates of hypermutation in different regions along the HIV-1 genome are indeed independent from one another. As discussed in section one of this thesis, Suspene *et al.* suggested that G to A hypermutation varied across the HIV-1 genome due to the time it remained ss during viral DNA replication (189). Another study conducted by Yu *et al.* observed varying levels of G residue changes in different regions of the HIV-1 genome. They observed that G to A nucleotide changes were the lowest and highest at the 5' and 3' end of the genome, respectively (211). This was in agreement with studies conducted by Suspene *et al.*(189) that gradients were established in the HIV-1 genome and it has been suggested that the gradients observed for regions of hypermutation were the result of evolution (211). Our data suggests that hypermutation is absent in some while detectable in other regions of HIV-1, independent of position or

gradient along the genome; however these findings are likely much more complicated due to other biological factors involved.

There are caveats to our data. We were unable to obtain exact sample dates from all subjects that had previous sequence analysis for vpu/env hypermutation. The sample dates for some subjects that had vpu/env proviral hypermutated DNA was different compared to the patient samples examined at the time of the study. However all samples were collected +/- 2 years from original sampling date. Out of a total 13 hypermutated vpu/env proviral DNA subjects from previous study, we were able to date-matched samples for half of these subjects due to limited sample availability. Also, generating PCR amplicons from some samples of subjects that were obtained for the present study (but could not be used) had undergone several cycles of freeze-thaw because of use in other studies. This would affect the stability of DNA and hence, may have altered the integrity of the study (therefore they were not used for conducting proviral hypermutation analysis). This further decreased our sample size. Due to factors such as small sample size, inconsistent patient samples, and reliability of DNA, these studies need to be repeated and validated. Future experiments of this nature would be to conduct sequencing data on better quality DNA so we could generate and also increase the sample size of the study.

5.4 Increased host APOBEC3G expression in *ex vivo* PBMCs of HIV⁺ women: Vif-mediated regulation

We were able to show that women with hypermutations in vpu/env/pol had more detectable endogenous APOBEC3G protein expression than women with non-

hypermutated provirus. However, endogenous Vif was expressed in most subjects from both groups at varying levels. This made it difficult to come to any conclusion on the role of Vif/APOBEC interaction based on the examination between the two groups of patients. Based on the chi-square analysis, we were unable to determine if a relationship persisted as there was a small sample size. If the study were to be conducted on a larger sample pool, then I would believe that subjects with increased proviral hypermutation in these regions have increased APOBEC3G.

Peripheral blood lymphocytes and monocytes can be strongly protected against HIV-1 infection as these cell types abundantly express APOBEC3G (91). In results section 4.2, Western blot analysis showed varying levels of APOBEC3G and Vif protein expression in frozen PBMCs of subjects with hypermutated and non-hypermutated proviruses. *Ex vivo* studies such as these are sometimes difficult to evaluate as there are several pathogen and host biological variables that exist and these variables must be well controlled for before coming to concrete conclusions (ie. Vif expression). We can hypothesize that increased APOBEC-mediated hypermutation could be due to either changes in the host or changes in the virus. For example, the overall replication of the virus (ie. number of circulating viral RNAs), viral subtype, and Vif variation all participate in overall pathogenesis in the host and hence will play a role in the number and frequency of APOBEC3G molecules present in a host and thus hypermutation levels.

Theoretically, subjects infected with non-hypermutated proviruses should demonstrate decreased APOBEC3G protein expression and increased Vif protein expression. One factor that plays a role in this expression is the amount of virus produced in an individual and this will obviously influence the phenotypic expression of proteins

(146). Pace *et al.* conducted near full-length population analysis and observed that viral loads from patients with hypermutated proviral DNA sequences were lower than those with non-hypermutated proviral DNA sequences (4.32 ± 0.60 versus 4.98 ± 0.75 \log_{10} copies HIV RNA/mL; $P = 0.001$) (146). However, a similar analysis by Land *et al.* did not find any significance ($p = 0.82$) in viral load data among hypermutated sequences (118). However, the viral load data presented from this group was dubious due to the utilization of frozen archived samples, but as earlier discussed, the data generated for CD4⁺ T cell counts ($p = 0.009$) and the data presented in this thesis strongly supports and suggests that women from this cohort may have the ability to maintain higher CD4⁺ T cell counts due to proviral hypermutation.

Ten out of twelve subjects with hypermutated proviruses in this study had detectable levels of endogenous APOBEC3G protein expression, while in only three out of eight subjects infected with non-hypermutated provirus was APOBEC3G expression observed. Our findings suggest that increased hypermutation seems to be associated with increased APOBEC3G expression in PBMCs. Janini *et al.* have shown that during PBMCs activation modification experiments, they were able to able to modify the hypermutation levels in PBMCs (95). Their proposed model showed that hypermutation by APOBEC3G activity was predominant in resting PBMCs, whereas APOBEC3F activity predominated in activated PBMCs (95). (95). A study conducted in September 2010 by Iwabu *et al.* clearly demonstrated that sequence diversity (based on subtype?) of the *vif* gene may influence APOBEC3G levels in PBMCs (91). From Figure 16, non-hypermutated sequences except for ML2713 and ML2767 are from subtype A1. ML2713 and ML2767 are from subtypes C and D, respectively which supports the idea that

subtype-based sequence differences in Vif may influence APOBEC3G protein levels. This leaves for discussion to determine if Vif proteins from different subtypes have varying levels of regulation on APOBEC3G proteins.

Recognizing subtype differences in relation to disease progression is of global importance. Kanki *et al.* were the first to observe that progression to AIDS differed between subtypes (104). This led to several studies showing that disease progression is affected by subtype differences in HIV-1 and these subtype differences are found in different genes of the virus (89, 104, 109). Most studies have not examined subtype differences in vif and the antiviral role of APOBEC3G. Iwabu *et al.*, was the first group to study biological differences, in terms of viral fitness and quasispecies, of Vif *in vivo* (91). From their data, they concluded that subtype C HIV-1 viruses are most protected against endogenous APOBEC-mediated hypermutation (91). (91). An earlier study supports these conclusions where subtype C viruses were highly protected from APOBEC3G-mediated hypermutation (95). Sixty percent of all HIV-1 infections worldwide are subtype C (89). With the combined knowledge this suggests that if the majority of circulating viruses are of subtype C, then it would be difficult to target the virus with an “APOBEC” therapy to these individuals due to the viral C subtype being fairly resistant to G to A hypermutation. Different vif clades may play a role in Vif/APOBEC interactions, however further studies must be conducted to characterize the relationship between subtype variation and antiviral activities of APOBEC3 proteins.

The results generated in section 4.3, lead us to consider our suggestion that different subtypes of Vif in this study population are regulating APOBEC3G expression differently. One of the difficulties associated with *ex vivo* studies in humans is that data

from real patients has some uncontrolled and some unmeasured variables such as vif sequence variations and hence, it is difficult to draw conclusions at this juncture.

It would be interesting to re-blot these subjects with an anti- APOBEC3F antibody as we know that both APOBEC3G and APOBEC3F are expressed at different levels due to factors such a cell type, and HIV disease status (97, 98, 149, 172). Due to time constraints and unavailability of this antibody, we were unable to test this. The reasoning for detecting APOBEC3F is that the subjects with hypermutated proviruses, who had low levels of APOBEC3G, may express higher expression of APOBEC3F since both proteins are known to cause APOBEC-mediated hypermutation. Based on the data presented in this thesis (from Table 6 and Figure 20) and previous data from our group by Dr. Allison Land (Hypermur 2.0 data not presented in this thesis), we should be able to examine data as we know which subject has increased levels of endogenous APOBEC3G and endogenous APOBEC3F activity. Examination of the hypermutation activity of both data sets in APOBEC3G and APOBEC3F appear and are consistent with the Western blot data from Figure 20.

Another factor to consider is to conduct this study on fresh PBMCs as opposed to frozen cells. Frozen cells are known to lose their viability when improperly frozen thru inappropriate handling, and repetitive freeze-thaw cycles when other samples are pulled. All these problems can affect the number of viable cells to work with. Further examination is required on a larger subset of subjects infected by different HIV-1 subtypes and their vif sequences must be fully characterized to conduct studies between Vif and APOBEC3G/F function. Since this study did reveal interesting results on a very small scale, these results must be replicated to enhance the final conclusions.

5.5 The relationship between the level of HIV-1 Vif expression and the level of host APOBEC3G: Vif-mediated regulation

To determine if Vif plays a role in APOBEC3G levels and hypermutation in this sample set, we examined expression of Vif from subjects with HIV-1 proviral sequences that are hypermutated or non-hypermutated. Vif from three subjects with increased proviral hypermutation and Vif from three non-hypermutated virus clones were cloned and expressed into a SCMV-IN mammalian expression vector. The SCMV-IN expression vectors containing different viral vif inserts revealed Vif protein expression differences among subjects (two subjects had undetectable expression when compared to the other four subjects and the positive control). These observed patterns of Vif expression were seen in all three experimental replicates and always showed the same effect (ie. expression of Vif from subject ML1418 was always absent or very low). The co-transfection of the vif plasmids with wt APOBEC3G (pCMV-HA) indicated that there was also striking variability observed in APOBEC3G protein expression. What makes Vif expression differ between subjects and does wt APOBEC3G degradation differ between hosts due to variation in Vif expression levels, or due to other Vif-dependent or – independent factors? It is important to note that all vectors used in this study (SCMV-IN, pcDNA-HVif, and pCMV-HA) were under the activity of the CMV promoter and they were expressed in the same mammalian cell line. Thus based on the results in section 4.4, it seems likely that the sequence differences between different HIV-1 strains may affect Vif expression.

Analysis of each vif sequence from all subjects did not demonstrate significant

APOBEC-mediated hypermutation at *vif*; however multiple sequence differences were identified at different regions throughout the *vif* genome (Figure 21, section 4.4). This suggests that sequence changes (other than APOBEC-mediated hypermutation) may well be affecting the transcription, translation, and expression of Vif. This may be an explanation as to why some Vif proteins were under expressed in whole cell lysates of transfected HEK 293 T cells when compared to each other and to the wt Vif control. *In vitro* expression of mammalian and viral proteins can be challenging as there are several reasons why they would not be fully expressed such as codon optimization, gene encoded transcription, translation, and post translational modifications, all of which have an effect on protein expression.

Sequence changes in genes can affect the transcript, translation and the post translation of Vif (or any protein) expression. The *vif* sequence changes from our subjects may have an impact on the outcome of HIV-1 disease as a study by Yamada *et al.* showed that the proviruses from long-term non-progressors had several mutations present in the accessory genes (where *vif* had the highest proviral mutation rate out of all the accessory genes), while the majority of progressors did not have any mutations in this region (209). This further supports the possibility that *vif* sequence variation (at the transcription level) from the six subjects may play a role in disease progression by initiating a sequence of events where the translation of mRNA may essentially affect the process all the way down to post translational modifications of Vif. And begs the question – how long have your subjects been infected? Further work is necessary to prove this conclusion as each step of the central dogma of molecular biology (DNA to RNA to protein) must be thoroughly examined. The DNA sequences and protein coding residues as previously

shown in Figure 21 were thoroughly examined; however it would be interesting to conduct further protein structure and functional studies on these Vif proteins to determine if there are differences in protein structure or function which would have an effect on the APOBEC3G protein interaction.

Several studies have noted that out of the four accessory proteins in HIV-1, Vif is unique due to it being the only protein that is rapidly degraded intracellularly (60, 199). The reason for Vif proteins to rapidly turn-over may be due to prevent itself from inducing detrimental effects such as suppressing its own viral infectivity (1) at high protein concentrations (60). As discussed, there are critical amino acid motifs required for Vif to enhance infectivity in cells as well as regions critical for binding APOBEC3 proteins. Several studies have characterized Vif function through mutational and deletion analysis experiments with Δ Vif mutants (45, 60, 94, 199). Some important amino acids involved in Vif function are at positions 14-17, 40-44, 63-70, and 86-89 (45, 60, 199). Interestingly, in all of our subjects with significant hypermutation at vpu/env (ML1592, ML1970, and ML1975), amino acids 63-70 have more than one amino acid change compared to the consensus sequence and their Vif expression was reduced compared to the wt control. The amino acid positions 63-70 (involved in formation of β -strand structures) are critical for normal expression of Vif proteins as well as infectivity enhancement. Another important region, 40-44, is specific to binding APOBEC3G (45). All three subjects with hypermutated proviruses had an amino acid change from phenylalanine to tyrosine at position 44. This may explain why subjects with these mutant vif viruses were not able to restrict APOBEC3G levels as well as wt Vif.

Vif undergoes degradation through a similar pathway as APOBEC3 proteins- the

ubiquitin-proteasome pathway. Fujiti *et al.* showed that Vif expression is increased by blocking this pathway. Perhaps the Vif constructs in our study have sequence variation making them more likely to be degraded through the ubiquitin-proteasomal pathway. Recent studies have shown that HIV-1 Vif is involved in cell cycling by inducing cell cycle disturbance in the G2 phase (45, 170, 200, 201). G2, the final subphase of interphase in the cell cycle, is the stage where rapid cell growth and protein synthesis occurs before a cell enters mitosis. A key component of cell cycle delay is the interaction of Vif with the ubiquitin ligase complex and it is important to note that disruption of the cell cycle is independent of specific APOBEC3 proteins such as APOBEC3D/E, F, and G (45). DeHart *et al.* illustrated important APOBEC3G/F binding sites located in Vif that are required for G2 accumulation (positions 14-17 and 40-44). These positions are required for Vif to interact to the Cul5 and Elongin B/C complex and undergo degradation of APOBEC3F and APOBEC3G, respectively (45). Our subjects with hypermutated proviruses all were shown to have at least one amino acid change within positions 40-44, suggesting that G2 accumulation should decrease (enter G2 arrest/delay) which means that cell replication and protein expression has curtailed or stopped. However, these G2 delay speculations must be further examined and confirmed by monitoring the cell cycle profiles of infected cells.

It's difficult to draw any clear conclusion as to why we observed such low Vif expression in our studies. One experimental limitation may be due to sequence variations in vif, as this may affect the transfection efficiency of vif uptake into 293 T cells and hence, affect the protein expression of wt APOBEC3G and Vif proteins in co-transfected cells. There are many sequence differences between the different vif sequences in our

subjects and in comparison between subjects; none of these Vif mutations are consistent. Furthermore, these amino acid changes in Vif could be attributed to G2 delay of the cell cycle, a key component in Vif expression. In this section, we have only discussed one small portion of how Vif induces G2 cell cycle delay. There are other studies that have also noted other amino acids changes in Vif that also play a role in G2 cycle arrest (94). Based on our vif sequencing data, it is possible that the amino acid changes in Vif result in a variety of changes affecting Vif expression and function. Further studies are required to understand the biological significance of the stability and G2 cycle delay by Vif.

It is difficult to examine APOBEC degradation with inconsistent vif expression as measured by Western blot and to come to a conclusion. Further examination and detailed experiments are required to thoroughly investigate the role of Vif in APOBEC degradation and HIV disease progression.

6.0 Conclusions and future directions

The work completed in this thesis has demonstrated the importance of the interplay of host and viral factors in protection against HIV-1. One of the main findings of this thesis is that the detected APOBEC3G and APOBEC3F changes in *pol* and *vpu/env* regions of the HIV-1 genome indicate independent activity of one another. Also, subjects with significant proviral hypermutation had higher expression of endogenous APOBEC3G. The characterization of subjects who have detectable hypermutated proviral DNA in the *pol* gene of HIV-1 and a positive correlation to CD4⁺ T cell counts, suggests an association between APOBEC-mediated hypermutation and disease progression. Vif expression differs in subjects and this variation makes it difficult to establish if Vif affects APOBEC3G degradation. Furthermore, protein expression differences in the host (APOBEC3G) and in the virus (Vif) suggest an influential role in HIV-1 disease progression. Additional studies are required to provide more evidence of APOBEC-mediated hypermutation and HIV-1 disease progression.

It is unfortunate that viral load data was unavailable for this study since it is a powerful determinant in HIV-1 disease progression status; it would be worthwhile to collect viral load data from subjects in future studies to better monitor HIV-1 disease progression. Further studies to characterize the role of APOBEC in HIV-1 disease progression should include the measurement mRNA APOBEC3G levels by real-time PCR as this could illustrate a difference in expression levels between the two groups of interest – hypermutated and non-hypermutated proviruses. Also, sequencing the APOBEC3G gene from these patients would locate single nucleotide polymorphisms (as

previously found from subjects with hypermutated proviruses at vpu/env (117, 118)) as could further explain disease progression differences within our cohort and other cohorts worldwide.

It would be beneficial to take information from the protein expression studies from our *ex vivo* and *in vivo* experiments and further investigate a larger population; this would help strengthen our conclusions. In addition to immunoblotting for APOBEC3G proteins from our whole cell lysates, it would be useful to acquire an APOBEC3F antibody to observe the protein expression composition differences between the two antiretroviral proteins.

Based on the above suggestions and difficulties observed in this thesis, an ideal study would investigate the best markers of disease progression through prospective studies (CD4⁺ T cell expression and viral load data), measure APOBEC3G/F protein expression, mRNA expression and hypermutation activity (sequence pol or vpu/env or ideally the whole virus), and finally further characterize Vif sequences and function from patients with and without proviral hypermutation.

In conclusion, APOBEC3 proteins serve as an important antiretroviral mechanism in hosts by inducing hypermutation along the HIV-1 genome. Understanding the mechanisms of these proteins and their relationship with HIV-1 Vif function will help the development of future novel antiretroviral therapies.

7.0 References

1. **Akari, H., M. Fujita, S. Kao, M.A. Khan, M. Shehu-Xhilaga, A. Adachi, and K. Strebel.** 2004. High level expression of human immunodeficiency virus type-1 Vif inhibits viral infectivity by modulating proteolytic processing of the Gag precursor at the p2/nucleocapsid processing site. *J. Biol. Chem.* **279**:12355-12362.
2. **Alce, T. M., and W. Popik.** 2004. APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. *J. Biol. Chem.* **279**:34083-34086.
3. **Alimonti, J. B., T.B. Ball, and K.R. Fowke.** 2003. Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS. *J. Gen. Virol* **7**:1649-1661.
4. **Ammann, A. J., G. Schiffman, D. Abrams, P. Volberding, J. Ziegler, and M. Conant.** 1984. B-cell immunodeficiency in acquired immune deficiency syndrome. *JAMA* **251**:1447-9.
5. **An, P., G. Bleiber, P. Duggal, G. Nelson, M. May, B. Mangeat, I. Alobwede, D. Trono, D. Vlahov, S. Donfield, J.J. Goedert, J. Phair, S. Buchbinder, S.J. O'Brien, A. Telenti, and C.A. Winkler.** 2004. APOBEC3G genetic variants and their influence on the progression to AIDS. *J. Virol.* **78**:11070-11076.
6. **Andersen, J. L., E. Le Rouzic, and V. Planelles V.** 2008. HIV-1 Vpr: mechanisms of G2 arrest and apoptosis. *Exp. Mol. Pathol.* **85**:2-10.
7. **Archary, D., M.L. Gordon, T.N. Green, H.M. Coovadia, P.J.R. Goulder, and T. Ndung'u.** 2010. HIV-1 subtype C envelope characteristics associated with divergent rates of chronic disease progression. *Retrovirology* **7**:92-104.
8. **Arhel, N.** 2010. Revisiting HIV-1 uncoating. *Retrovirology.* **7**:1-10.
9. **Armitage, A. E., A. Katzourakis, T. de Oliveira T J.J. Welch, R. Belshaw, K.N. Bishop, B. Kramer, A.J. McMichael, A. Rambaut, A.K. Iversen.** 2008. Conserved footprints of APOBEC3G on Hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences. *J. Virol.* **82**:8743-8761.
10. **Baeten, J. M., B. Chohan, L. Lavreys, V. Chohan, R.S. McClelland, L. Certain, K. Mandaliya, W. Jacoko, and J. Overbaugh.** 2007. HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads. *J. Infect. Dis.* **195**:1177-1180.
11. **Bergeron, J. R. C., H. Huthoff, D.A. Veselkov, R.L. Beavil, P.J. Simpson, S.J. Matthews, M.H. Malim, and M.R. Sanderson.** 2010. The SOCS-Box of HIV-1 Vif interacts with elonginBC by induced-folding to recruit its cul5-containing ubiquitin ligase complex. *PLoS Pathog* **6**:e1000925.
12. **Bishop, K. N., R.K. Holmes, A.M. Sheehy, N. O. Davidson, S.J. Cho, and M.H. Malim.** 2004. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr. Biol.* **14**:1392-1396.
13. **Brenchley, J. M., D.A. Price, and D.C. Douek.** 2006. HIV disease: fallout from a mucosal catastrophe? *Nat. Immunol.* **7**:235-239.

14. **Broadhead, R. L.** 1996. Transmission of HIV-1 in breast milk. *Trop. Doct* **3**:100-101.
15. **Buchbinder, S. P., D.V. Mehrotra, A. Duerr, D.W. Fitzgerald, R. Mogg, D. Li, P.B. Gilbert, J.R. Lama, M. Marmor, C. Del Rio, M.J. McElrath, D.R. Casimiro, K.M. Gottesdiener, J.A. Chodakewitz, L. Corey, M.N. Robertson, and Step Study Protocol Team.** 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet*:1881-1893.
16. **Bukrinskaya, A., B. Brichacek, A. Mann, and M. Stevenson.** 1998. Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. *J. Exp. Med.* **188**:2113-2125.
17. **Cadima-Couto, I. a. J. G.** 2010. Towards inhibition of Vif-APOBEC3G interaction: which protein to target? *Advanc. Virol.*
18. **Campbell, E. M., M.P. Dodding, M.W. Yap, X. Wu, S. Gallois-Montbrun, M.H. Malim, J.P. Stoye, and T.J. Hope.** 2007. TRIM5 alpha cytoplasmic bodies are highly dynamic structures. *Mol. Biol. Cell* **18**:2102-2111.
19. **Cen, S., F. Guo, M. Niu, J. Saadatmand, J. Deflassieux, and L. Kleiman.** 2004. The interaction between HIV-1 Gag and APOBEC3G. *J. Biol. Chem.* **279**:33177-33184.
20. **Cen, S., Z.G. Peng, X.Y. Li, Z.R. Li, J. Ma, Y.M. Wang, B. Fan, X.F. You, Y.P. Wang, F. Liu, R.G. Shao, L.X. Zhao, L. Yu, J.D. Jiang.** 2010. Small molecular compounds inhibit HIV-1 replication through specifically stabilizing APOBEC3G. *J. Biol. Chem.* **285**:16546-16552.
21. **Chaisson, R. E., J.E. Gallant, J.C. Keruly, and R.D. Moore.** 1998. Impact of opportunistic disease on survival in patients with HIV infection. *AIDS* **1**:29-33.
22. **Chan, D. C., and P.S. Kim.** 1998. HIV entry and its inhibition. *Cell* **93**:681-684.
23. **Charneau, P., G. Mirambeau, P. Roux, S. Paulous, H. Buc, and F. Clavel.** 1994. HIV-1 reverse transcription. A termination step at the center of the genome. *J. Mol. Biol.* **241**:651-662.
24. **Chen, B., E.M. Vogan, H. Gong, J.J. Skehel, D.C. Wiley, and S.C. Harrison.** 2005. Determining the structure of an unliganded and fully glycosylated SIV gp120 envelope glycoprotein. *Structure* **2**:197-211.
25. **Chen, Z., T.L. Eggerman, A.V. Bocharov, I.N. Baranova, T.G. Vishnyakova, G. Csako, A.P. Patterson.** 2010. Hypermutation induced by APOBEC-1 overexpression can be eliminated. *RNA* **16**:1040-1052.
26. **Cho, S. J., H. Drechsler, R.C. Burke, M.Q. Arens, W. Powderly, and N.O. Davidson.** 2006. APOBEC3F and APOBEC3G mRNA levels do not correlate with human immunodeficiency virus type 1 plasma viremia or CD4+ T-cell count. *J. Virol.* **80**:2069-2072.
27. **Chomont, N., M. El-Far, P. Ancuta, L. Trautmann, F.A. Procopio, B. Yassine-Diab, G. Boucher, M.R. Boulassel, G. Ghattas, J.M. Brenchley, T.W. Schacker, B.J. Hill, D.C. Douek, J.P. Routy, E.K. Haddad, and R.P. Sekaly.** 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med* **8**:893-900.

28. **Cohen, C. R., R. Nguti, E.A. Bukusi, H. Lu, C. Shen, M. Luo, S. Sinei, F. Plummer, J. Bwayo, and R.C. Brunham.** 2000. Human immunodeficiency virus type 1-infected women exhibit reduced interferon-gamma secretion after *Chlamydia trachomatis* stimulation of peripheral blood lymphocytes. *J. Infect. Dis.* **182**:1672-1677.
29. **Cohen, M. S.** 2006. Amplified transmission of HIV-1: missing link in the HIV pandemic. *Trans. Am. Clin. Climatol. Assoc.* **117**:213-225.
30. **Cohen, M. S., N. Hellmann, J.A. Levy, K. DeCock, and J. Lange.** 2008. The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. *J. Clin. Invest.* **118**:1244-1254.
31. **Connor, R. I., H. Mohri, Y. Cao, and D.D. Ho.** 1993. Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals. *J. Virol.* **67**:1772-1777.
32. **Conticello, S. G.** 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* **9**:229.
33. **Conticello, S. G., C.J.F. Thomas, S.K. Petersen-Mahrt, and M.S. Neuberger.** 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases. *Mol. Biol. Evol.* **22**:367-377.
34. **Copeland, K. F. T.** 2006. Inhibition of HIV-1 entry into cells. *Recent Pat. Antiinfect. Drug Discov.* **1**:107-112.
35. **Daelemans, D., E. Afonina, J. Nilsson, G. Werner, J. Kjems, E. De Clercq, G.N. Pavlakis, and A.M. Vandamme.** 2002. A synthetic HIV-1 Rev inhibitor interfering with the CRM1-mediated nuclear export. *Proc. Natl. Acad. Sci.* **99**:14440-14445.
36. **Darby, S. C., D.W. Ewart, P.L. Giangrande, R.J. Spooner, and C.R., Rizza.** 1996. Importance of age at infection with HIV-1 for survival and development of AIDS in UK haemophilia population. UK Haemophilia Centre Director's Organisation. *Lancet* **347**:1573-1579.
37. **Darbyshire, J.** 2000. Therapeutic interventions in HIV infection - a critical view. *Trop. Med. Int. Health* **5**:A26-A31.
38. **Datta, P., J.E. Embree, J.K. Kreiss, J.O. Ndinya-Achola, M. Braddick, M. Temmerman, N.J.D. Nagelkerke, G. Maitha, K.K. Holmes, P. Piot, H.O. Pamba, F.A. Plummer.** 1994. Mother-to-child transmission of human immunodeficiency virus type 1: report from the Nairobi study. *J. Infect. Dis.* **5**:1134-1140.
39. **de Arellano, E. R., J. Alcamí, M. López, V. Soriano, and A. Holguín.** 2010. Drastic decrease of transcription activity due to hypermutated long terminal repeat (LTR) region in different HIV-1 subtypes and recombinants. *Antiviral Res.* **88**:152-159.
40. **de Bethune, M. P.** 2010. Non-nucleoside reverse transcriptase inhibitors (NNRTIs), their discovery, development, and use in the treatment of HIV-1 infection: a review of the last 20 years (1989-2009). *Antiviral Res.* **1**:75-90.
41. **De Clercq, E.** 2009. Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *Int. J. Antimicrob. Agents* **4**:307-320.

42. **De Clercq, E.** 2009. The history of antiretrovirals: key discoveries over the past 25 years. *Rev. Med. Virol.* **5**:287-299.
43. **de Wolf, F., I. Spijkerman, P.T. Schellekens, M. Langendam, C. Kuiken, M. Bakker, M. Roos, R. Coutinho, F. Miedema, and J. Goudsmit.** 1997. AIDS prognosis based on HIV-1 RNA, CD4+T cell count and function: markers with reciprocal predictive value over time after seroconversion. *AIDS* **15**:1799-1806.
44. **Deforche, K., T. Silander, R. Camacho, Z. Grossman, M.A. Soares, K. Van Laethem, R. Kantor, Y. Moreau, A.M. Vandamme, and non-B Workgroup.** 2006. Analysis of HIV-1 pol sequence using Bayesian networks: implications for drug resistance. *Bioinformatics* **24**:2975-2979.
45. **DeHart, J. L., A. Bosque, R.S. Harris, and V. Planelles.** 2008. Human Immunodeficiency Virus Type 1 Vif Induces Cell Cycle Delay via Recruitment of the Same E3 Ubiquitin Ligase Complex That Targets APOBEC3 Proteins for Degradation. *J. Virol.* **82**:9265-9272.
46. **DeHart, J. L., E.S. Zimmerman, O. Ardon, C.M. Monteiro-Filho, E.R. Argañaraz, and V. Planelles.** 2007. HIV-1 Vpr activates the G2 checkpoint through manipulation of the ubiquitin proteasome system. *Virol. J.* **4**:57.
47. **Douglas, J. L., J. K. Gustin, K. Viswanathan, M. Mansouri, A. V. Moses, and K. Früh.** 2010. The Great Escape: Viral Strategies to Counter BST-2/Tetherin. *PLoS Pathog* **5**:e10000913.
48. **Dussupt, V., P. Sette, N.F. Bello, M.P. Javid, K. Nagashima, and F. Bouamr.** 2011. Basic Residues in the Nucleocapsid Domain of Gag Are Critical for Late Events of HIV-1 Budding. *J. Virol.* **85**:2304-2315.
49. **Easterbrook, P. J., M. Smith, J. Mullen, S. O'Shea, I. Chrystie, A. de Ruiter, I.D. Tatt, A.M. Geretti, and M. Zuckerman.** 2010. Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy. *J. Int. AIDS Soc.* **13**:4.
50. **Eckert, D. M., and P.S. Kim.** 2001. Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.* **70**:777-810.
51. **Esnault, C., O. Heidmann, F., Delebecque, M. Dewannieux, D. Ribet, A.J. Hance, T. Heidmann, and O. Schwartz.** 2005. APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature* **433**:430-433.
52. **Fiorentini, S., E. Marini, S. Caracciolo, and A. Caruso.** 2006. Functions of the HIV-1 matrix protein p17. *New Microbiol.* **29**:1-10.
53. **Foster, J. L., S.J. Denial, B.R. Temple, and J.V. Garcia.** 2011. Mechanisms of HIV-1 Nef Function and Intracellular Signaling. *J. Neuroimmune. Pharmacol.* **Epub ahead of print.**
54. **Fowke, K. R., N.J. Nagelkerke, J. Kimani, J.N. Simonsen, A.O. Anzala, J.J. Bwayo, K.S. MacDonald, E.N. Ngugi, and F.A. Plummer.** 1996. Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. *Lancet* **348**:1347-1351.
55. **Fowke, K. R., R. Kaul, K.L. Rosenthal, J. Oyugi J. Kimani, W.J. Rutherford, N.J. Nagelkerke, T.B. Ball, J.J. Bwayo, J.N. Simonsen, G.M. Shearer, and F.A. Plummer.** 2000. HIV-1-specific cellular immune responses among HIV-1 resistant sex workers. *Immunol. Cell Biol.* **78**:586-595.

56. **Freed, E. O.** 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* **251**:1-15.
57. **Freed, E. O.** 2001. HIV-1 replication. *Somat. Cell Mol. Genet.* **26**:13-33.
58. **Freed, E. O., and M.A. Martin.** 2001. HIVs and Their Replication, p. 1971-2041. In K. D. H. L. W. a. Wilkins (ed.), In *Fields Virology*, 4 ed, vol. 2.
59. **Fritz, J. V., D. Dujardin, J. Godet, P. Didier, J. De Mey, J.L. Darlix, Y. Mély, and H. de Rocquigny.** 2010. HIV-1 Vpr oligomerization but not that of Gag directs the interaction between Vpr and Gag. *J. Virol.* **84**:1585-1596.
60. **Fujita, M., H. Akari, A. Sakurai, A. Yoshida, T. Chiba, K. Tanaka, K. Strebel, and A. Adachi.** 2004. Expression of HIV-1 accessory protein Vif is controlled uniquely to be low and optimal by proteasome degradation. *Microbes. Infect* **6**:791-798.
61. **Gabuzda, D. A., K. Lawrence, E. Langhoff, E. Terwilliger, T. Dorfman, W.A. Haseltine, and J. Sodroski.** 1992. Role of *vif* in replication of human immunodeficiency virus type 1 in CD4⁺ T lymphocytes. *J. Virol.* **66**:6489-6495.
62. **Galai, N., A. Kalinkovich, R. Burstein, D. Vlahov, and Z. Bentwich.** 1997. African HIV-1 subtype C and rate of progression among Ethiopian immigrants in Israel. *Lancet* **349**:180-181.
63. **Gallo, R. C. a. M., L.** 2003. The discovery of HIV as the cause of AIDS. *N. Engl. J. Med.* **24**:2283-2285.
64. **Gandhi, S. K., J.D. Siliciano, J.R. Bailey, R.F. Siliciano, and J.N. Blankson.** 2008. Role of APOBEC3G/F-mediated hypermutation in the control of human immunodeficiency virus type 1 elite suppressors. *J. Virol.* **82**:3125-3130.
65. **Gao, F., L. Yue, A.T. White, P.G. Pappas, J. Barchue, A.P. Hanson, B.M. Greene, P.M. Sharp, G.M. Shaw, and B.H. Hahn.** 1992. Human infection by genetically diverse SIVSM-related HIV-2 in west Africa. *Nature* **358**:495-499.
66. **Gaschen, B., J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B.H. Hahn, T. Bhattacharya, and B. Korber.** 2002. Diversity considerations in HIV-1 vaccine selection. *Science* **296**:2354-2360.
67. **Geyer, M., O.T. Fackler, and B.M. Peterlin BM.** 2001. Structure--function relationships in HIV-1 Nef. *EMBO Rep.* **2**:580-585.
68. **Gill, C. J., J.L. Griffith, D. Jacobson, S. Skinner, S.L. Gorbach, and I.B. Wilson.** 2002. Relationship of HIV Viral Loads, CD4 Counts, and HAART Use to Health-Related Quality of Life. *J. Acquir. Immune Defic. Syndr.* **30**:485-492.
69. **Goldschmidt, V., G. Bleiber, M. May, R. Martinez, M. Ortiz, A. Telenti, and Swiss HIV Cohort Study.** 2006. Role of common human TRIM5alpha variants in HIV-1 disease progression. *Retrovirology* **3**:54.
70. **Goncalves, J. a. M. S.-M.** 2004. HIV-1 Vif and APOBEC3G: multiple roads to one goal. *Retrovirology* **1**:28-35.
71. **Gorry, P. R. a. P. A.** 2010. Coreceptors and HIV-1 pathogenesis. *Curr HIV/AIDS Rep.*
72. **Gottlinger, H. G.** 2001. HIV-1 gag: A molecular machine driving viral particle assembly and release. Theoretical Biology and Biophysics Group, Los Alamos, NM, LA-UR 02-2877.
73. **Greenblatt, R. M., S.A. Lukehart, F.A. Plummer, T.C. Quinn, C.W.**

- Critchlow, R.L. Ashley, L.J. D'Costa, J.O. Ndinya-Achola, L. Corey, A.R. Ronald, et al.** 1988. Genital ulceration as a risk factor for human immunodeficiency virus infection. *AIDS* **2**:47-50.
74. **Grosskurth, H., F. Plummer, F. Mhalu, and D. Mabey.** 1993. STD research in Africa. *Lancet* **342**:1415-1416.
75. **Han, Y., X. Wang, Y. Dang, Y. H. Zheng.** 2007. APOBEC3G and APOBEC3F Require an Endogenous Cofactor to Block HIV-1 Replication. *PLoS Pathog* **4**:e1950000.
76. **Hansasuta, P., and S.L. Rowland-Jones.** 2001. HIV-1 transmission and acute HIV-1 infection. *Br. Med. Bull.* **58**:109-127.
77. **Harris, R. S., K. Bishop, A.M. Sheehy, H.M. Craig, S. K. Petersen-Mahrt, I.N. Watt, M.S. Neuberger, and M.H. Malim.** 2003. DNA deamination mediates innate immunity to retroviral infection. *Cell* **6**:803-809.
78. **Harris, R. S. a. M. T. L.** 2004. Retroviral restriction by Apobec proteins. *Nat. Rev. Immunol.* **4**:868-877.
79. **Heeney, J. L., A.G. Dalgleish, and R.A. Weiss.** 2006. Origins of HIV and the evolution of resistance to AIDS. *Science* **313**:462-466.
80. **Hessol, N. A., L.A. Napolitano, D. Smith, Y. Lie, A. Levine, M. Young, M. Cohen, H. Minkoff, K. Anastos, G. D'Souza, R.M. Greenblatt, and J.J. Goedert.** 2010. HIV tropism and decreased risk of breast cancer. *PloS One* **5**:e14349.
81. **Hewson, T., N. Lone, M. Moore, and S. Howie.** 1999. Interactions of HIV-1 with antigen-presenting cells. *Immunol. Cell Biol.* **77**:289-303.
82. **Hoffman, T. L., and R.W. Doms.** 1999. HIV-1 envelope determinants for cell tropism and chemokine receptor use. *Mol. Membr. Biol.* **16**:57-65.
83. **Hogg, R. S., B. Yip, K.J. Chan, E. Wood, K.J. Craib, M.V. O'Shaughnessy, J.S. Montaner.** 2001. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA* **286**:2568-2577.
84. **Holmes, R. K., F.A. Koning, K.N. Bishop, and M.H. Malim.** 2007. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J. Biol. Chem.* **282**:2587-2595.
85. **Holmes, R. K., M.H. Malim, and K.N. Bishop.** 2007. APOBEC-mediated viral restriction: not simply editing? *Trends Biochem. Sci.* **32**:118-128.
86. **Hu, D. J., A. Buve, J. Baggs, G. van der Groen, and T.J. Dondero.** 1999. What role does HIV-1 subtype play in transmission and pathogenesis? An epidemiological perspective. *AIDS* **13**:873-881.
87. **Hughes, M. D., V.A. Johnson, M.S. Hirsch, J.W. Bremer, T. Elbeik, A. Erice, D.R. Kuritzkes, W.A. Scott, S.A. Spector, N. Basgoz, M.A. Fischl, R.T. D'Aquila.** 1997. Monitoring plasma HIV-1 RNA levels in addition to CD4+ lymphocyte count improves assessment of antiretroviral therapeutic response. ACTG 241 Protocol Virology Substudy Team. *Ann. Intern. Med.* **126**:983-985.
88. **Huthoff, H., and G.J. Towers.** 2008. Restriction of retroviral replication by APOBEC3G/F and TRIM5alpha. *Trends Microbiol.* **12**:612-619.
89. **Iordanskiy, S., M. Waltke, Y. Feng, and C. Wood.** 2010. Subtype-associated

differences in HIV-1 reverse transcription affect the viral replication. *Retrovirology* **7**:85.

90. **Isel, C., R. Marquet, G. Keith, C. Ehresmann, and B. Ehresmann.** 1993. Modified nucleotides of tRNA^{3^{Lys}} modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. *J. Biol. Chem.* **268**:25269-25272.
91. **Iwabu, Y., M. Kinomoto, M. Tatsumi, H. Fujita, M. Shimura, Y. Tanaka, Y. Ishizaka, D. Nolan, S. Mallal, T. Sata, and K. Tokunaga.** 2010. Differential anti-APOBEC3G activity of HIV-1 Vif proteins derived from different subtypes. *J. Biol. Chem.* **285**:35350-35358.
92. **Iwatani, Y., D.S. Chan, F. Wang, K.S. Maynard, W. Sugiura, A.M. Gronenborn, I. Rouzina, M.C. Williams, K. Musier-Forsyth, and J.G. Levin.** 2007. Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res.* **21**:7096-7108.
93. **Iwatani, Y., D.S.B. Chan, L. Liu, H. Yoshii, J. Shibata, N. Yamamoto, J.G. Levin, A.M. Gronenborn, and W. Sugiura.** 2009. HIV-1 Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain. *Proc. Natl. Acad. Sci.* **106**:19539-19544.
94. **Izumi, T., K. Io, M. Matsui, K. Shirakawa, M. Shinohara, Y. Nagai, M. Kawahara, M. Kobayashi, H. Kondoh, N. Misawa, Y. Koyanagi, T. Uchiyama, and A. Takaori-Kondo.** 2010. HIV-1 viral infectivity factor interacts with TP53 to induce G2 cell cycle arrest and positively regulate viral replication. *Proc. Natl. Acad. Sci.* **107**:20798-20803.
95. **Janini, M., M. Rogers, D.R. Birx, and F.E. McCutchan.** 2001. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J. Virol.* **75**:7973-7986.
96. **Jennes, W., S. Verheyden, C. Demanet, C.A. Adje-Toure, B. Vuylsteke, J.N. Nkengasong, and et al.** 2006. Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. *J. Immunol.* **177**:6588-6592.
97. **Jin, X., A. Brooks, H. Chen, R. Bennett, R. Reichman, and H. Smith.** 2005. APOBEC3G/CEM15 (hA3G) mRNA Levels Associate Inversely with Human Immunodeficiency Virus Viremia. *J. Virol.* **79**:11513-11516.
98. **Jin, X., H. Wu, and H. Smith.** 2007. APOBEC3G levels predict rates of progression to AIDS. *Retrovirology* **4**:20.
99. **John, G. C. a. J. K.** 1996. Mother-to-child transmission of human immunodeficiency virus type 1. *Epidemiol. Rev.* **2**:149-157.
100. **Joly, V. a. P. Y.** 1999. Non nucleoside reverse transcriptase inhibitors. *AIDS* **1**:37-44.
101. **Kahn, J. O. a. B. D. W.** 1998. Acute human immunodeficiency virus type 1 infection. *N. Engl. J. Med.* **339**:33-39.
102. **Kajaste-Rudnitski, A., C. Pultrone, F. Marzetta, S. Ghezzi, T. Coradin, and E. Vicenzi.** 2010. Restriction factors of retroviral replication: the example of

- Tripartite Motif (TRIM) protein 5 alpha and 22. *Amino Acids* **39**:1-9.
103. **Kaleebu, P., N. French, C. Mahe, D. Yirrell, C. Watera, F. Lyagoba, J. Nakiyingi, A. Rutebemberwa, D. Morgan, J. Weber, C. Gilks, and J. Whitworth.** 2002. Effect of human immunodeficiency virus (HIV) type 1 envelope subtypes A and D on disease progression in a large cohort of HIV-1 positive persons in Uganda. *J. Infect. Dis.* **9**:1244-1450.
 104. **Kanki, P. J., D.J. Hamel, J.L. Sankale, C.C. Hsieh, I. Thior, F. Barin, S. A. Woodcock, A. Gueye-Ndiaye, E. Zhang, M. Montano, T. Siby, R. Marlink, I. NDoeye, M. E. Essex, and S. MBoup.** 1999. Human immunodeficiency virus type 1 subtypes differ in disease progression. *J. Infect. Dis.* **179**:68-73.
 105. **Kaul, R., D. Trabattoni, J.J. Bwayo, D. Arienti, A. Zagliani, F. M. Mwangi, C. Kariuki, E.N. Ngugi, K.S. MacDonald, T.B. Ball, M. Clerici, and F.A. Plummer.** 1999. HIV-1-specific mucosal IgA in a cohort of HIV-1-resistant Kenyan sex workers. *AIDS* **13**:23-29.
 106. **Kieffer, T. L., P. Kwon, R.E. Nettles, Y. Han, S.C. Ray, and R.F. Siliciano.** 2005. G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J. Virol.* **79**:1975-1980.
 107. **Kijak, G. H., L.M. Janini, S. Tovanabuttra, E. Sanders-Buell, M.A. Arroyo, M.L. Robb, N.L. Michael, D.L. Birx, and F.E. McCutchan.** 2008. Variable contexts and levels of hypermutation in HIV-1 proviral genomes recovered from primary peripheral blood mononuclear cells. *Virology* **376**:101-111.
 108. **Killen, J., C. Grady, G.K. Folkers, and A.S. Fauci.** 2002. Ethics of clinical research in the developing world. *Nat. Rev. Immunol.* **2**:210-215.
 109. **Kiwanuka, N., O. Laeyendecker, M. Robb, G. Kigozi, M. Arroyo, F. McCutchan, L.A. Eller, M. Eller, F. Makumbi, D. Birx, F. Wabwire-Mangen, D. Serwadda, N.K. Sewankambo, T.C. Quinn, M. Wawer, and R. Gray.** 2008. Effect of human immunodeficiency virus type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J. Infect. Dis.* **197**:707-713.
 110. **Kleinman, L., C.P. Jones, and K. Musier-Forsyth.** 2010. Formation of the tRNA^{Lys} packaging complex in HIV-1. *FEBS Lett.* **2**:359-365.
 111. **Knight, S. C.** 1994. A problem of antigen presentation? *Curr. Biol.* **12**:1131-1134.
 112. **Knoepfel, S. A., F. Di Giallonardo, M. Däumer, A. Thielen, and K.J. Metzner.** 2011. In-depth analysis of G-to-A hypermutation rate in HIV-1 env DNA induced by endogenous APOBEC3 proteins using massively parallel sequencing. *J. Virol. Methods* **171**:329-338.
 113. **Krebs, F. C., T.H. Hogan, S. Quiterio, S. Gartner, B. Wigdahl.** 2001. Lentiviral LTR-directed Expression, Sequence Variation, and Disease Pathogenesis. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
 114. **Kuhl, B. D., V. Cheng, M.A. Wainberg, and C. Liang.** 2011. Tetherin and Its Viral Antagonists. *J. Neuroimmune. Pharmacol.* **Epub ahead of print.**
 115. **Lama, J., A. Mangasarian, and Trono D.** 1999. Cell-surface expression of CD4

- reduces HIV-1 infectivity by blocking Env incorporation in a Nef- and Vpu-inhibitable manner. *Curr. Biol.* **9**:622-631.
116. **Lanao, J. M., E. Briones, C.I. Colino.** 2007. Recent advances in delivery systems for anti-HIV1 therapy. *J. Drug Target* **1**:21-36.
 117. **Land, A. M.** 2009. The role of HIV-1 recombination and APOBEC3G/F-mediated hypermutation in HIV-1 pathogenesis. University of Manitoba, Winnipeg.
 118. **Land, A. M., T.B. Ball, M. Luo, R. Pilon, P. Sandstrom, J.E. Embree, C. Wachih, J. Kimani, F.A.Plummer.** 2008. Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J. Virol.* **82**:8172-8182.
 119. **Lauring, A. S., and R. Andino.** 2010. Quasispecies theory and the behaviour of RNA viruses. *PLoS Pathog* **7**:1001005.
 120. **Leavitt, A. D., L. Shiue, and H.E. Varmus.** 1992. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J. Biol. Chem.* **268**:2113-2119.
 121. **Lecossier, D., F. Bouchonnet, F. Clavel, and A.J. Hance.** 2003. Hypermutation of HIV-1DNA in the absence of the Vif protein. *Science* **300**:1112.
 122. **Lee, H. L., A.S. Perelson, S.C. Park, and T. Leitner.** 2008. Dynamic correlation between intrahost HIV-1 quasispecies evolution and disease progression. *PLoS Comput. Biol.* **12**:e1000240.
 123. **listed, N. a.** 1998. HIV-1 viral load and CD4 cell count in untreated children with vertically acquired asymptomatic or mild disease. Paediatric European Network for Treatment of AIDS (PENTA). *AIDS* **12**:F1-8.
 124. **Liu, S., W. Jing, B. Cheung, H. Lu, J. Sun, X. Yan, J. Niu, J. Farmer, S. Wu, and S. Jiang.** 2007. HIV gp41 C-terminal heptad repeat contains multifunctional domains. Relation to mechanisms of action of anti-HIV peptides. *J. Biol. Chem.* **282**:9612-9620.
 125. **Liu, Y., F. Li, Y. Liu, K. Hong, X. Meng, J. Chen, Z. Zhang, Z. Huo, M. Sun, S.G. Self, and Y. Shao.** 2011. HIV fragment gag vaccine induces broader T cell response in mice. *Vaccine* [Epub ahead of print].
 126. **Liu, Y., M.R. Nonnemacher, D.L. Stauff, L. Li, A. Banerjee, B. Irish, E. Kilaeski, N. Rajagopalan, J.B. Suchitra, Z.K. Khan, U. Ranga, and B. Wigdahl.** 2010. Structural and functional studies of CCAAT/enhancer binding sites within the human immunodeficiency virus type 1 subtype C LTR. *Biomed. Pharmacother.* **10**:672-680.
 127. **Luo, K., T.Wang, B. Liu, C. Tian, Z. Xiao, J.Kappes, and X.F.** 2007. Cytidine Deaminases APOBEC3G and APOBEC3F Interact with Human Immunodeficiency Virus Type 1 Integrase and Inhibit Proviral DNA Formation. *J. Virol.* **81**:7238-7248.
 128. **Mangeat, B., P.Turelli, G. Caron, M. Friedli, L. Perrin, and D. Trono.** 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**:99-103.
 129. **Martin, D. J.** 2004. Immunology of HIV infection and the host response. *In* T. S. Marlink RG (ed.), *From the Ground Up: Building Comprehensive HIV/AIDS*

- Care Programs in Resource-Limited Settings, vol. 2. Elizabeth Glaser Pediatric AIDS Foundation; 2009, Washington, DC.
130. **Mbisa, J. L., W. Bu, V.K. Pathak.** 2010. APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. *J. Virol.* **10**:5250-5259.
 131. **McCune, J. M.** 2001. The dynamics of CD4+ T-cell depletion in HIV disease. *Nature* **410**:974-979.
 132. **McMichael, A. J. a. T. H.** 2003. HIV vaccines 1983-2003. *Nat Med* **9**:874-880.
 133. **Mercenne, G., S. Bernacchi, D. Richer, G. Bec, S. Henriet, J.C. Paillart, and R. Marquet.** 2009. HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation. *Nucleic Acids Res.* **2**:633-646.
 134. **Moore, J. P., F.E. McCutchan, S.W. Poon, J. Mascola, J. Liu, Y. Cao and D.D. Ho.** 2004. Exploration of antigenic variation in gp120 from clades A through F of human immunodeficiency virus type 1 by using monoclonal antibodies. *J. Virol.* **68**:8350-8364.
 135. **Muesing, M. A., D.H. Smith, C.D. Cabradilla, C.V. Benton, L.A. Lasky, D.J. Capon.** 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature* **313**:450-458.
 136. **Mukadi, Y. D., D. Maher, and A. Harries.** 2001. Tuberculosis case fatality rates in high HIV prevalence populations in sub-Saharan Africa. *AIDS* **15**:143-152.
 137. **Nakayama, E. E., and T. Shioda.** 2010. Anti-retroviral activity of TRIM5 alpha. *Rev. Med. Virol.* **20**:77-92.
 138. **Nannini, E. C. a. P. C. O.** 2002. HIV1 and the gut in the era of highly active antiretroviral therapy. *Curr. Gastroenterol. Rep.* **5**:392-398.
 139. **Nasio, J. M., N.J. Nagelkerke, A. Mwatha, S. Moses, J.O. Ndinya-Achola, and F.A. Plummer.** 1996. Genital ulcer disease among STD clinic attenders in Nairobi: association with HIV-1 and circumcision status. *Int. J. STD AIDS* **7**:410-414.
 140. **Newman, E. N., R.K. Holmes, H.M. Craig, K.C. Klein, J.R. Lingappa, M.H. Malim, and A.M. Sheehy.** 2005. Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. *Curr. Biol.* **15**:166-170.
 141. **Ngugi, E. N., and F.A. Plummer.** 1988. Health outreach and control of HIV infection in Kenya. *J. Acquir. Immune Defic. Syndr.* **1**:566-570.
 142. **Niewiadomska, A. M., and X.F. Yu.** 2009. Host restriction of HIV-1 by APOBEC3 and viral evasion through vif. *Curr. Top. Microbiol. Immunol.* **339**:1-25.
 143. **O'Brien, S. J., and G.W. Nelson.** 2004. Human genes that limit AIDS. *Nat. Genet.* **36**:565-574.
 144. **O'Brien, S. J., P.H. Hartigan, D. Martin, J. Esinhart, A. Hill, S. Benoit, M. Rubin, M.S. Simberkoff, J.D. Hamilton, and the Veterans Affairs Cooperative Study Group on AIDS.** 1996. Changes in plasma HIV-1 RNA and CD4+ T lymphocyte counts and the risk of progression to AIDS. *N. Engl. J. Med.* **334**:426-431.
 145. **Ogg, G. S., X. Jin, S. Bonhoeffer, P.R. Dunbar, M.A. Nowak, S. Monard, J.P. Segal, Y. Cao, S.L. Rowland-Jones, V. Cerundolo, A. Hurley, M. Markowitz, D.D. Ho, D.F. Nixon, and A.J. McMichael.** 1998. Quantitation of HIV-1-

- Specific Cytotoxic T Lymphocytes and Plasma Load of Viral RNA. *Science* **279**:2103-2106.
146. **Pace, C., J. Keller, D. Nolan, I. James, S. Gaudieri, C. Moore, and S. Mallal.** 2006. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and *vif* genetic variation. *J. Virol.* **80**:9259-9269.
 147. **Pantaleo, G., C. Graziosi, and A. Fauci.** 1993. The immunopathogenesis of human immunodeficiency virus infection. *N. Engl. J. Med.* **328**:327-335.
 148. **Paraschiv, S., B. Foley, and D. Otelea.** 2010. Diversity of HIV-1 subtype C strains isolated in Romania. *Infect. Infect. Genet. Evol.* [Epub ahead of print].
 149. **Peng, G., K.J. Lei, W. Jin, T. Greenwell-Wild, and S.M. Wahl.** 2006. Induction of APOBEC3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV-1 activity. *J. Exp. Med.* **203**:41-46.
 150. **(PEPFAR), U. S. P. s. E. P. f. A. R.** 2011, posting date. Treatment. [Online.]
 151. **Piantadosi, A., D. Humes, B. Chohan, R.S. McClelland, and J. Overbaugh.** 2009. Analysis of the percentage of human immunodeficiency virus type 1 sequences that are hypermutated and markers of disease progression in a longitudinal cohort, including one individual with a partially defective *vif*. *J. Virol.* **16**:7805-7814.
 152. **Pilcher, C. D., D.C. Shugars, S.A. Fiscus, W.C. Miller, P. Menezes, J. Giner, B. Dean, K. Robertson, C.E. Hart, J.L. Lennox, J.J. Eron, and C.B. Hicks.** 2001. HIV in body fluids during primary HIV infection: implications for pathogenesis, treatment, and public health. *AIDS* **15**:837-845.
 153. **Piot, P., M. Bartos, P. D. Ghys, N. Walker, and B. Schwartlander.** 2001. The global impact of HIV/AIDS. *Nature* **410**:968-973.
 154. **Plantier, J. C., M. Leoz, J.E. Dickerson, F. De Oliveira, F. Cordonnier, V. Leme, F. Damond, D.L. Robertson, and F. Simon.** 2009. A new human immunodeficiency virus derived from gorillas. *Nat. Med.* **8**:871-872.
 155. **Plummer, F. A., T.B. Ball, J. Kimani, and Fowke KR.** 1999. Resistance to HIV-1 infection among highly exposed sex workers in Nairobi: what mediates protection and why does it develop? *Immunol. Lett.* **66**:27-34.
 156. **Poignard, P., E.O. Saphire, P.W. Parren, and D.R. Burton.** 2001. gp120: Biologic aspects of structural features. *Annu. Rev. Immunol.* **19**:253-274.
 157. **Prochnow, C., R. Bransteitter, and X.S. Chen.** 2009. APOBEC deaminases-mutases with defensive roles for immunity. *Sci. China C. Life Sci.* **52**:893-902.
 158. **Qu, W., M. Robinson, and F.J. Zhang.** 2008. Factors influencing the natural history of HIV-1 infection. *Chin. Med. J.* **121**:2613-2621.
 159. **Raney, A., A.Y. Shaw, J.L. Foster, and J.V. Garcia.** 2007. Structural constraints on human immunodeficiency virus type 1 Nef function. *Virology* **368**:7-16.
 160. **Rayne, F., S. Debaisieux, A. Bonhoure, and B. Beaumelle.** 2010. HIV-1 Tat is unconventionally secreted through the plasma membrane. *Cell Biol. Int.* **34**:409-413.
 161. **Re, M. C., F. Vitone, C. Biagetti, P. Schiavone, F. Alessandrini, I. Bon, E. de Crignis, D. Gibellini.** 2010. HIV-1 DNA proviral load in treated and untreated

- HIV-1 seropositive patients. *Clin. Microbiol. Infect.* **16**:640-646.
162. **Reddy, K., C.A. Winkler, L. Werner, K. Milsana, S.S. Abdool Karim, T. Ndung'u, and CAPRISA Acute Infection Study Team.** 2010. APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. *AIDS* **24**:195-204.
163. **Reyes-Darias, J. A., F.J. Sanchez-Luque, and A. Berzal-Herranz.** 2008. Inhibition of HIV-1 replication by RNA-based strategies. *Curr. HIV Research* **6**:500-514.
164. **Romani, B., S. Engelbrecht, and R.H. Glashoff.** 2010. Functions of Tat: the versatile protein of human immunodeficiency virus type 1. *J. Gen. Virol* **91**:1-12.
165. **Rose, P. P., and B.T. Korber.** 2000. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* **16**:400-401.
166. **Rosenberg, E. S., J.M. Billingsley, A.M. Caliendo, S.L. Boswell, P.E. Sax, S.A. Kalams, and B.D. Walker.** 1997. Vigorous HIV-1-Specific CD4 T cell responses associated with control of viremia. *Science* **278**:1447-1450.
167. **Ross, T. M.** 2001. Using death to one's advantage: HIV modulation of apoptosis. *Leukemia* **3**:332-341.
168. **Royce, R. A., A. Sena, W. Cates, and M.S. Cohen.** 1997. Sexual transmission of HIV. *N. Engl. J. Med.* **336**:1072-1078.
169. **Rudnicka, D. a. O. S.** 2009. Intrusive HIV-1-infected cells. *Nat. Immunol.* **10**:933-934.
170. **Sakai, K., J. Dimas, and M.J. Lenardo.** 2006. The Vif and Vpr accessory proteins independently cause HIV-1-induced T cell cytopathicity and cell cycle arrest. *Proc. Natl. Acad. Sci.* **103**:3369-3374.
171. **Santiago, M. L., F. Range, B.F. Keele, Y. Li, E. Bailes, F. Bibollet-Ruche, C. Fruteau, R. Noë, M. Peeters, J.F. Brookfield, G.M. Shaw, P.M. Sharp, and B.H. Hahn.** 2005. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (*Cercocebus atys atys*) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *J. Virol.* **79**:12515-12527.
172. **Sarkis, P. T., S. Ying, R. Xu, and X.F. Yu.** 2006. STAT1-independent cell type-specific regulation of antiviral APOBEC3G by IFN- α . *J. Immunol.* **177**:4530-4540.
173. **Sasada, A., A. Takaori-Kondo, K. Shirakawa, M. Kobayashi, A. Abudu, M. Hishizawa, K. Imada, Y. Tanaka, and T. Uchiyama.** 2005. APOBEC3G targets human T-cell leukemia virus type 1. *Retrovirology* **2**:32.
174. **Scarlata, S., and C. Carter.** 2003. Role of HIV-1 Gag domains in viral assembly. *Biochim. Biophys. Acta.* **1614**:62-72.
175. **Schwartz, R. A.** 2004. Kaposi's sarcoma. An update. *J. Surg. Oncol.* **87**:146-151.
176. **Schwartz, S., M. Campbell, G. Nasioulas, J. Harrison, B.K. Felber, and G.N. Pavlakis.** 1992. Mutational inactivation of an inhibitory sequence in human immunodeficiency virus type 1 results in Rev-independent gag expression. *J. Virol.* **12**:7176-7182.
177. **Sen, S., S. Mandal, S. Bhattacharya, S. Halder, and P. Bhaumik.** 2010. Oral manifestations in human immunodeficiency virus infected patients. *Indian J.*

- Dermatol. **1**:116-118.
178. **Shao, Q., Y. Wang, J.E.K. Hildreth, and B. Liu.** 2010. Polyubiquitination of APOBEC3G is essential for its degradation by HIV-1 Vif. *J. Virol.* **84**:4840-4844.
 179. **Sharma, Y., U. Neogi, V. Sood, S. Banerjee, S. Samrat, A. Wanchu, S. Singh, and A.C. Banerjee.** 2010. Genetic and functional analysis of HIV-1 Rev Responsive Element (RRE) sequences from North-India. *AIDS. Res. Ther.* **7**:28.
 180. **Shojania, S., and J.D. O'Neil.** 2010. Intrinsic disorder and function of the HIV-1 Tat protein. *Protein Pept. Lett* **17**:999-1011.
 181. **Smith, H. C.** 2011. APOBEC3G: a double agent in defense. *Trends Biochem. Sci.* **Epub ahead of print.**
 182. **Sorg, T. a. M. M.** 1997. Gene therapy for AIDS. *Transfus. Sci.* **18**:277-289.
 183. **Stebbing, J. a. G. M.** 2003. The clades of HIV: their origins and clinical significance. *AIDS Rev.* **4**:205-213.
 184. **Stevenson, M.** 2003. HIV-1 pathogenesis. *Nat. Med.* **7**:853-860.
 185. **Stieler, K., and N. Fischer.** 2010. Apobec 3G efficiently reduced infectivity of the human exogenous gammaretrovirus XMRV. *PloS One* **5**:e11738.
 186. **Strebel, K.** 2005. APOBEC3G & HTLV-1: Inhibition without deamination. *Retrovirology* **2**:37.
 187. **Strebel, K.** 2003. Virus-host interactions: role of HIV proteins Vif, Tat, and Rev. *AIDS* **17**:S25-S34.
 188. **Strebel, K., D. Daugherty, K. Clouse, D. Cohen, T. Folks, and M.A. Martin.** 1987. The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature* **328**:728-730.
 189. **Suspene, R., C. Rusniok, J.P. Vartanian, and S. Wain-Hobson.** 2006. Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic Acids Res.* **34**:4677-4684.
 190. **Tomaras, G. D., N.L. Yates, P. Liu, L. Qin, G.G. Fouda, L.L. Chavez, A.C. Decamp, R. J. Parks, V.C. Ashley, J.T. Lucas, M. Cohen, J. Eron, C.B. Hicks, H.X.Liao, S.G. Self, G. Landucci, D.N. Forthal, K.J. Weinhold, B.F. Keele, B.H. Hahn, M.L. Greenberg, L. Morris, S.S.A. Karim, W.A. Blattner, D.C. Montefiori, G.M. Shaw, A.S. Perelson, and B.F. Haynes.** 2008. Initial B-Cell Responses to Transmitted Human Immunodeficiency Virus Type 1: Virion-Binding Immunoglobulin M (IgM) and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia. *J. Virol.* **82**:12449-12463.
 191. **Ulunga, N. K., A.D. Sarr, D. Hamel, J.L. Sankale, S. Mboup, and P.J. Kani.** 2008. The level of APOBEC3G (hA3G) - related G-to-A mutations does not correlate with viral load in HIV type 1-infected individuals. *AIDS Res. Hum. Retroviruses* **24**:1285-1290.
 192. **(UNAIDS), J. U. N. P. o. H. A.** 2011. Global Report: UNAIDS report on the global AIDS epidemic 2010.
 193. **Vaccari, M., P. Poonam, and G. Franchini.** 2010. Phase III HIV vaccine trial in Thailand: a step toward a protective vaccine for HIV. *Expert Rev. Vaccines* **9**:997-1005.
 194. **Van Damme, N., D. Goff, C. Katsura, R.L. Jorgenson, R. Mitchell, M.C.**

- Johnson , E.B. Stephens, and J. Guatelli J.** 2008. The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe* **3**:245-252.
195. **Vartanian, J. P., D. Guetard, M. Henry, and S. Wain-Hobson.** 2008. Evidence of editing human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* **320**:230-233.
196. **Vázquez-Pérez, J. A., C.E. Ormsby, R. Hernández-Juan, K.J. Torres, and G. Reyes-Terán.** 2009. APOBEC3G mRNA expression in exposed seronegative and early stage HIV infected individuals decreases with removal of exposure and with disease progression. *Retrovirology* **6**:23.
197. **Vidal, F., F. Gutierrez, M. Gutierrez, M. Olona, V. Sanchez, G. Mateo, J. Peraire, C. Vilades, S. Veloso, M. Lopez-Dupla, and P. Domingo.** 2010. Pharmacogenetics of adverse effects due to antiretroviral drugs. *AIDS Rev.* **12**:15-30.
198. **Volpe, F., A. Schwimmer, and C. Barr.** 1985. Oral manifestation of disseminated *Mycobacterium avium intracellulare* in a patient with AIDS. *Oral Surg. Oral Med. Oral Pathol.* **6**:567-570.
199. **Wang, H., A. Sakurai, B. Khamsri, T. Uchiyama, H. Gu , A. Adachi, and M. Fujita M.** 2005. Unique characteristics of HIV-1 Vif expression. *Microbes. Infect* **7**:385-390.
200. **Wang, J., E.L. Reuschel, J. M. Shackelford, L. Jeang, D.K. Shivers, J. A. Diehl, X.F. Yu, and T. H. Finkel.** 2011. HIV-1 Vif promotes the G1- to S-phase cell-cycle transition. *Blood* **117**:1260-1290.
201. **Wang, J., J.M. Shackelford, C.R. Casella, D.K. Shivers, E.L. Rapaport, B. Liu, X.F. Yu, and T.H. Finkel.** 2007. The Vif accessory protein alters the cell cycle of human immunodeficiency virus type 1 infected cells. *Virology* **359**:243-252.
202. **Wang, J., J.M. Shackelford, N. Selliah, D.K. Shivers , E. O'Neill, J.V. Garcia, K. Muthumani, D. Weiner, X.F. Yu, D. Gabuzda, T.H. Finkel.** 2008. The HIV-1 Vif protein mediates degradation of Vpr and reduces Vpr-induced cell cycle arrest. *DNA Cell. Biol.* **27**:267-277.
203. **Wang, X., P.T. Dolan, Y. Dang, and Y.H. Zheng.** 2007. Biochemical differentiation of APOBEC3F and APOBEC3G proteins associated with HIV-1 life cycle. *J. Biol. Chem.* **282**:1585-1594.
204. **Wang, Y., Y. Yang, Z. Hu, X. Chen, and G. Chen.** 2008. A systems biology analysis of protein-protein interactions in the APOBEC family. *Life Sci.* **83**:521-530.
205. **Weidner, J., U. Cassens, W. Göhde, W. Sibrowski, G. Odaibo, D. Olaleye, D. Reichelt, B. Greve.** 2010. An improved PCR method for detection of HIV-1 proviral DNA of a wide range of subtypes and recombinant forms circulating globally. *J. Virol. Methods* **172**:22-26.
206. **Wissing, S., N.L. Galloway, and W.C. Greene.** 2010. HIV-1 Vif versus the APOBEC3 cytidine deaminases: an intracellular duel between pathogen and host restriction factors. *Mol. Aspects Med* **5**:383-397.
207. **Wyatt, R. a. J. S.** 1998. The HIV-1 envelope glycoproteins: fusogens, antigens,

- and immunogens. *Science* **280**:1884-1888.
208. **Xu, H., E. Chertova, J. Chen, D.E. Ott, J.D. Roser, W.S. Hu, and V.K. Pathak.** 2007. Stoichiometry of the antiviral protein APOBEC3G in HIV-1 virions. *Virology* **360**:247-256.
209. **Yamada, T., and A. Iwamoto A.** 2000. Comparison of proviral accessory genes between long-term nonprogressors and progressors of human immunodeficiency virus type 1 infection. *Arch. Virol.* **145**:1021-1027.
210. **Yeni, P. G., S.M. Hammer, C.C. Carpenter, D.A. Cooper, M.A. Fischl, J.M. Gatell, B.G. Gazzard, M.S. Hirsch, D.M. Jacobsen, D.A. Katzenstein, J.S. Montaner, D.D. Richman, M.S. Saag, M. Schechter, R.T. Schooley, M.A. Thompson, S. Vella, and P.A. Volberding.** 2002. Antiretroviral treatment for HIV adult infection in 2002: update recommendations of the International AIDS Society-USA Panel. *JAMA* **2**:222-235.
211. **Yu, Q., R. König, S. Pillai, K. Chiles, M. Kearney, S. Palmer, D. Richman, J.M. Coffin, and N.R. Landau.** 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat. Struct. Mol. Biol.* **11**:435-442.
212. **Zennou, V., and P.D. Bieniasz.** 2006. Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. *Virology* **349**:31-40.
213. **Zennou, V., C. Petit, D. Guetard, U. Nerhbass, L. Montagnier, and P. Charneau.** 2000. HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell* **101**:173-185.
214. **Zhang, H., B. Yang, R.J. Pomerantz, C. Zhang, S.C. Arunachalam, and L. Gao.** 2003. The cytidine demethylase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**:94-98.
215. **Zhang, H., R. J. Pomerantz, G. Dornadula, and Y. Sun.** 2000. Human Immunodeficiency Virus Type 1 Vif Protein Is an Integral Component of an mRNP Complex of Viral RNA and Could Be Involved in the Viral RNA Folding and Packaging Process. *J. Virol.* **74**.
216. **Zhang, R., R. Mehla, and A. Chauhan.** 2010. Perturbation of host nuclear membrane component RanBP2 impairs the nuclear import of human Immunodeficiency Virus-1 Preintegration Complex (DNA). *PloS One* **5**:1-13.
217. **Zhen, A., T. Wang, K. Zhao, Y. Xiong, and X.F. Yu.** 2010. A single amino acid difference in human APOBEC3H variants determines HIV-1 Vif sensitivity. *J. Virol.* **84**:1902-1911.
218. **Zhujun, A., Z. Yu, L. Wang, Y. Zheng, and X. Yao.** 2008. Vpr14-88-Apobec3G Fusion Protein Is Efficiently Incorporated into Vif-Positive HIV-1 Particles and Inhibits Viral Infection. *PloS One* **3**:e1995.

8.0 Abbreviations used

A	adenine
AID	activation-induced deaminase
AIDS	acquired immune deficiency syndrome
APOBEC	apolipoprotein B mRNA editing enzyme, catalytic polypeptide
APOBEC3	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
APOBEC3G	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G
att	attachment
AZT	zidovudine
BCA	Bicinchoninic acid
BST-2	bone marrow stromal antigen 2
C	cytidine
Cat. No.	catalogue number
CD4	cluster of differentiation 4
CD8	cluster of differentiation 8
CRF	circulating recombinant form
CSW	commercial sex worker
ddH ₂ O	double-distilled water
°C	degrees Celsius
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates

ds	double-stranded
env	envelope
EtBr	ethidium bromide
FBS	fetal bovine serum
G	guanine
g	gram
gag	group specific antigen
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
HAART	highly active antiretroviral therapy
HIV	human immunodeficiency virus
HIV-1	human immunodeficiency virus type 1
HIV-2	human immunodeficiency virus type 2
HIV ⁺	HIV-1 positive
HR	heptad region
HTLV-1	human T-cell leukemia virus type 1
IN	integrase
kDa	kilodalton
L	litre
LB	luria-bertani
LTR	long-terminal repeat
M	major
mM	millimolar
MCS	multiple cloning site

mg	milligram
mL	millilitre
mRNA	messenger ribonucleic acid
N	non-M and non-O
nef	negative regulator factor
NML	National Microbiology Laboratory
nm	nanometer
NNRTI	non-nucleoside reverse transcriptase inhibitor
NRTI	nucleoside analogue reverse transcriptase inhibitor
O	outlier
ori	origin of replication
p7	nucleocapsid
p17	matrix protein
p24	capsid protein
(PAK2)	p21 activated protein kinase
PBMC	peripheral blood mononuclear cell
pbs	primer binding site
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PEPFAR	United States President's Emergency Plan for AIDS Relief
PIC	preintegration complex
PI	protease inhibitor
pol	polymerase

PolyA	polyadenylation
PR	protease
PPT	polypurine tract
psi	packaging element
R	repeated
RRE	Rev Responsive Element
rev	differential regulator of expression
RIP 3.0	Recombination Identification Program 3.0
RIPA	radioimmunoprecipitation lysis buffer
RNA	ribonucleic acid
RNaseH	ribonuclease H
RPMI	Roswell park memorial institute medium
RTC	reverse transcription complex
RT	reverse transcriptase
SDS	sodium dodecyl sulfate
SIV	simian immunodeficiency virus
ss	single-stranded
T	thymine
tat	trans-activator of transcription
TE	tris EDTA
TBE	tris-borate EDTA
TBS	tris buffered saline
Trim5- α	tripartite motif 5-alpha

tRNA	transfer RNA
U	uracil
U3'	unique 3' sequence
U5'	unique 5' sequence
μl	microlitre
μM	micromolar
UNAIDS	Joint United Nations Programmes on HIV/AIDS
vif	viral infectivity factor
vpr	viral protein R
vpu	virus protein U
wt	wild-type